Compendium of Indicators for Evaluating Reproductive Health Programs

Volume 2

MEASURE
Evaluation

USAID

Part III
Indicators for Specific Programmatic Areas

A. **Global Reproductive Health Indicators**

B. **Family Planning and Fertility**

C. **STI/HIV/AIDS**

D. **Safe Motherhood**

E. **Newborn Health**

F. **Women's Nutrition**

G. **Breastfeeding**

H. **Adolescent Reproductive Health Programs**

I. **Postabortion Care (PAC)**

J. **Male Involvement in Reproductive Health Programs**

K. **Violence Against Women (VAW)**

L. **Female Genital Cutting (FGC)**

M. **Reproductive Health in Emergency Situations**

INDICATORS FOR SPECIFIC PROGRAMMATIC AREAS

Part III of the *Compendium* covers 12 programmatic areas for reproductive health, beginning with the 3 that command the largest percentage of government and donor budgets: **family planning, STI/HIV/AIDS,** and **safe motherhood**. Additional areas include **women's nutrition, newborn health**, and **breastfeeding**. During the 1990s governments and/or NGOs developed programmatic initiatives to meet a broader range of RH needs, including **adolescent reproductive health programs, postabortion care (PAC), male involvement, violence against women (VAW), female genital cutting (FGC),** and **reproductive health in emergency situations.**

Part III begins with two sets of RH indicators intended to measure RH status in countries worldwide. The first represents the response of the World Health Organization (WHO) in collaboration with selected reproductive health experts to monitor the extent to which programs/countries achieve progress toward the ICPD goals. The short list consists of 17 indicators that measure progress in a particular area (e.g., contraceptive prevalence rate, availability of basic essential obstetric care, HIV prevalence in pregnant women).

The second set of indicators – developed by Population Action International (PAI) – is the Reproductive Risk Index, which rates countries on 10 RH measures (many of which overlap with the WHO global monitoring indicators).

A primary difference between the two sets of indicators is that the WHO does not combine indicators into a summary score for each country, whereas the Reproductive Risk Index does. The latter raises some methodological concerns but is useful for advocacy purposes in that it yields comparative ratings for developing countries worldwide.

Part III of the *Compendium* outlines the main results that different RH programs are designed to achieve. Most indicators in Part III measure outputs or outcomes, although a few relate to the subjects covered in Part II (e.g., policy environment, quality of care). The exceptions tend to occur (1) where programs are relatively new and experts feel the need to stress policy and quality of care issues, or (2) where measures of behavior among the general public are not readily available. Although we include several long term measures of health status (e.g., fertility, mortality), the indicators in this section constitute the most frequently used measures for evaluating behavior change and impact (e.g., summative evaluation) for reproductive health programs.

Part III.A
Global Reproductive Health Indicators

- WHO's Short List of Reproductive Health Indicators for Global Monitoring
- PAI's Reproductive Risk Index

WHO's Short List of Reproductive Health Indicators for Global Monitoring

Definition
These 17 largely population-based indicators provide an overview of the reproductive health situation at the global and national level, endorsed by the WHO and the United Nations Interagency Working Group. (See Box III.A.1 for the listing of the 17 indicators and definitions of each.)

Data Requirements
Information on each of the 17 indicators (Note: Most of the indicators in Box III.A.1 are described in full elsewhere in this *Compendium*)

Data Source(s)
The DHS or other representative surveys of the intended population can provide certain indicators (1, 2, 4, 5, 8, 11, 13, 14, 15, and 17 in Box III.A.1). Other indicators (6, 7, 10, 12 and 16) require program-level data: service statistics, facility-based services, or laboratory results on clients. Whereas data are generally available for indicators based on the DHS or RHS, data may be difficult to obtain for certain measures (e.g., percentage of OB-GYN admissions owing to abortion).

Purpose and Issues
During the 1990s, the official representatives of countries worldwide attended international conferences (ICPD in Cairo, the Fourth Women's Conference in Beijing) and endorsed a number of global goals and targets in the broad area of sexual and reproductive health. This endorsement led to a proliferation of reproductive health indicators on which countries were asked to report. Subsequently, the United Nations asked the WHO to take the lead in organizing an interagency technical process to examine the issue of reproductive health indicators and to reach consensus on a short list of indicators for global monitoring.

The resulting set of 17 indicators covers the main areas of reproductive health and represents the consensus among international agencies of the key indicators for international comparison, global monitoring, and follow-up to the international conferences.

The purpose of this set of indicators is to provide an overview of the reproductive health situation at global and national levels. The objective is not to present a comprehensive set of indicators for program monitoring and evaluation. However, the data collected for reporting the indicators should be useful at the program management level.

The WHO has started to compile estimates for all 17 indicators. This set of indicators is not meant to serve as an index; rather, it draws attention to the key measurable areas of reproductive health.

**Box III.A.1 Definitions of WHO's Short List of Reproductive Health Indicators
for Global Monitoring (WHO, 2000a, 2001a)**

1. **Total Fertility Rate (TFR)**
    Total number of children a woman would have by the end of her reproductive period if she experienced the currently prevailing age-specific fertility rates throughout her childbearing life

2. **Contraceptive Prevalence Rate (CPR)[1]**
    Percent of women of reproductive age (15-49) who are using (or whose partner is using) a contraceptive method at a particular point in time

3. **Maternal Mortality Ratio (MMR)**
    Annual number of maternal deaths per 100,000 live births

4. **Antenatal Care Coverage**
    Percent of women attended at least once during pregnancy, by skilled health personnel (excluding trained or untrained traditional birth attendants), for reasons relating to pregnancy

5. **Percent of Births Attended by Skilled Health Personnel**
    Percent of births attended by skilled health personnel (excluding trained or untrained traditional birth attendants)

6. **Availability of Basic Essential Obstetric Care**
    Number of facilities with functioning basic essential obstetric care per 500,000 population

7. **Availability of Comprehensive Essential Obstetric Care**
    Number of facilities with functioning comprehensive essential obstetric care per 500,000 population

8. **Perinatal Mortality Rate (PMR)**
    Number of perinatal deaths per 1,000 total births

9. **Low Birth Weight Prevalence**
    Percent of live births that weigh less than 2,500g

10. **Positive Syphilis Serology Prevalence in Pregnant Women**
    Percent of pregnant women (15-24) attending antenatal clinics, whose blood has been screened for syphilis, with positive serology for syphilis

11. **Prevalence of Anemia in Women**
    Percent of women of reproductive age (15-49) screened for hemoglobin levels with levels 110g/l for pregnant women, and 120g/l for non-pregnant women

---

[1] The expert group working with WHO on this set of indicators recommends basing the calculation of contraceptive prevalence on all women of reproductive age, in contrast to the convention used by the DHS and RHS to report it for married women only (or married and unmarried women separately).

12. Percent of Obstetric and Gynecological Admissions Owing to Abortion
Percent of all cases admitted to service delivery points providing in-patient obstetric and gynecological services, which are due to abortion (spontaneous and induced, but excluding planned termination of pregnancy)

13. Reported Prevalence of Women with FGC
Percent of women interviewed in a community survey reporting having undergone FGC

14. Prevalence of Infertility in Women
Percent of women of reproductive age (15-49) at risk of pregnancy (not pregnant, sexually active, non-contracepting, and non-lactating) who report trying for a pregnancy for two years or more

15. Reported Incidence of Urethritis in Men
Percent of men aged (15-49) interviewed in a community survey reporting episodes of urethritis in the last 12 months

16. HIV Prevalence among Pregnant Women
Percent of pregnant women (15-24) attending antenatal clinics, whose blood has been screened for HIV and who are sero-positive for HIV

17. Knowledge of HIV-related Prevention Practices
Percent of all respondents who correctly identify all three major ways of preventing the sexual transmission of HIV and who reject three major misconceptions about HIV transmission or prevention

PAI's Reproductive Risk Index

Definition

Population Action International's (PAI) "report card" of nations composed of 10 key indicators of reproductive health. (See Box III.A.2 of key indicators.) Indicators used for the Reproductive Risk Index are each scored on a 100-point scale, and the scores are averaged to yield a total country score on the index. Based on their overall scores, countries are classified by their risk level: very high, high, moderate, low, and very low.

Data Requirements

Information on the ten key indicators

Data Source(s)

The data for each of the indicators come from diverse sources (DHS, WHO-estimates of MMR, and others)

Purpose and Issues

PAI's Reproductive Risk Index measures the progress of nations toward achieving the goals set at the International Conference on Population and Development in 1994. This tool is useful for advocacy purposes to document the poor state of sexual and reproductive health in much of the developing world. A total of 133 countries were ranked, representing 95 percent of the world's population.

PAI's Reproductive Risk Index includes many of the same indicators as the WHO's short list of reproductive health indicators for global monitoring. In fact, the two have seven indicators in common. In contrast to the WHO set which are not "combined" into an index, PAI has used scores on individual indicators to produce a summary score. This approach is open to methodological criticism that it implicitly gives equal weight to all ten indicators and "combines apples and oranges." However, the summary score allows for a classification of countries by level of risk, which is useful for advocacy purposes. The five risk categories are as follows:

Very high risk (60 points or more): The 19 countries in this category are characterized by early and high fertility, together with limited care during pregnancy and childbirth, all of which contribute to extremely high levels of maternal mortality. Safe and legal abortion is all but unavailable and, with few exceptions, levels of HIV infection are significant. All of the countries in this category have low average incomes; all but three are in sub-Saharan Africa.

High risk (45 – 59 points): The 26 countries in this category generally have low levels of contraceptive use, restrictive abortion policies, high birth rates, and high maternal mortality. HIV prevalence varies widely. Seventeen high risk countries are in sub-Saharan Africa, while the remaining nine are among the poorest nations in their respective regions.

Moderate risk (30 – 44 points): Women from these 28 countries have, on average, fewer than five children. Obtaining a safe and legal abortion is difficult or impossible in most countries in this category. The 28 countries in this category represent all the developing regions.

Low risk (15 – 29 points): In most of these 35 countries, fewer than one in twenty teenage girls gives birth annually and women have, on average, fewer than three children. Abortion is available on request in many of these countries. HIV prevalence is below one percent of adults in all but one of the countries in this category.

Very low risk (less than 15 points): Women in these 25 countries bear, on average, two or fewer children. Contraceptive use is high while anemia among pregnant women, HIV prevalence, and deaths from pregnancy and childbirth are low. Abortion is available on request in nearly all countries in this category. Fully 21 of the 25 countries in this category are wealthy, industrialized countries.

Box III.A.2 The Ten Indicators Used in PAI's Reproductive Risk Index:	Corresponding Section in this *Compendium*
1. Annual Births per 100 Women Aged 15-19	Part III.B[2]
2. Percent of Women Using Contraception	Part III.B
3. Abortion Policies	Part III.I
4. Prevalence of Anemia among Pregnant Women	Part III.F
5. Percent of Women Receiving Prenatal Care	Part III.D
6. Percent of Births Attended by Skilled Personnel	Part III.D
7. Percent of HIV/AIDS in Men	Part III.C
8. Percent of HIV/AIDS in Women	Part III.C[3]
9. Total Fertility Rate (TFR)	Part III.B
10. Maternal Deaths per 100,000 live births	Part III.D

[2] Indicator in Part III.B reads **Age Specific Fertility Rates (ASFR)**
[3] Indicator in Part III.C reads **HIV Prevalence Among Pregnant Women 15-24 years old**

Part III.B
Family Planning and Fertility

- Family Planning Program Effort Index

- Helping Individuals Achieve their Reproductive Intentions (HARI Index)

- Number of acceptors new to modern contraception

- Couple-years of protection (CYP)

- Method mix

- Contraceptive prevalence rate (CPR)

- Source of supply (by method)

- Continuation rates

- Unmet need for family planning

- Desire for additional children

- Wanted total fertility rate (WTFR)

- Age specific fertility rates (ASFR)

- Total fertility rate (TFR)

- Unwanted total fertility rate (UTFR)

Program evaluation is more methodologically advanced for family planning than for any other area of reproductive health, thanks to 30+ years of dedicated effort and strong funding support for this work. Several factors spurred the development of evaluation methodologies. The demographic concern that underscored the early family planning programs translated to monitoring of results in quantitative terms: numbers of new acceptors and continuing users. The problem of the "unequal weight" of different methods (e.g., one condom acceptor versus one vasectomy acceptor) gave rise to the index of couple-years of protection (CYP).

Critics and skeptics of family planning unwittingly strengthened evaluation efforts. The hot debate originating during the 1970s on the relative contribution of family planning programs versus that of socio-economic development impelled researchers to develop methods of demonstrating the **independent** effect of family planning. The family planning program effort scores (Lapham and Mauldin, 1984) played a useful role in this research. Although this public debate has subsided, the challenge remains to demonstrate causality (i.e., the program interventions have impact) in family planning program evaluation.

The World Fertility Survey (conducted from 1972 to 1984) focused primarily on the determinants of fertility, with relatively little attention to family planning and related programmatic issues. The Contraceptive Prevalence Survey, first piloted in 1975 in El Salvador, was designed to produce more programmatically useful results with shorter turn-around time. This survey tool later evolved into the Demographic and Health Survey (DHS) and the Reproductive Health Survey (RHS), both national-level representative surveys of women (and increasingly of men) of reproductive age in developing countries (Robey et al., 1992). The DHS and RHS have become the most widely utilized sources of data for family planning program evaluation, because of the quality of the information, standardization of items in the core questionnaire, availability of repeat measurement over time, and possibility of cross-national comparisons. Both surveys extend beyond family planning to cover related issues of maternal health/safe pregnancy and child health. In addition to the core questionnaire, the DHS offers 12 optional modules, including several of particular relevance to reproductive health (e.g., women's status and empowerment, violence against women, female genital cutting, HIV/AIDS, and maternal mortality). The RHS has a young adult version, and more recently in Africa, modules on sexual behavior and HIV testing and counseling.

A number of factors facilitate the job of evaluating family planning programs. First, the "intermediate outcome" – the desired behavioral change at the population level – is a single measurable behavior: use of a contraceptive method (aggregated to a measure of contraceptive prevalence). Second, despite the sensitive nature of family planning in many countries (especially in the early years of programs), women are willing and able to report contraceptive use with a high degree of accuracy, assuming the interviewer creates good rapport and the question is clearly worded. Third, the intermediate outcome of contraceptive prevalence is strongly (inversely) correlated with a key long-term outcome: fertility (except in countries with high levels of abortion). In short, family planning evaluators are blessed with a single, measurable, and valid outcome variable. Other areas of reproductive health (with the possible exception of breastfeeding) present a greater methodological challenge. Nonetheless, family planning program evaluation does have a few problems of its own.

Methodological Challenges of Evaluating Family Planning Programs

- **Contraceptive methods vary in terms of their use-effectiveness, duration of action, and likelihood of continuation.**

Although all "program methods"[1] can protect against pregnancy, some do so better than others. Thus, the program with a higher percentage of users of long-term methods will generally be more effective in pregnancy prevention than those in which users opt for less effective methods will be. The measure "couple-years of protection" (CYP) was designed to address this issue, but it has certain problems of its own (discussed below).

- **Large-scale survey data yield the most reliable estimates of contraceptive use, but they have limitations.**

Valuable as DHS/RHS data are for tracking national trends, they have three major limitations. First, such surveys are conducted only once every three to five years (if that often). Second, they do not yield precise results for geographical sub-areas (e.g., district, provinces) in most countries, which is the level at which program managers generally need their information. Third, these large-scale surveys are very expensive to conduct and analyze, a fact that has caused some countries to question the feasibility of continuing to fund them, especially if donor funding is unavailable. Given these limitations, program statistics, such as number of acceptors and CYP, are widely used to monitor family planning programs on a routine basis.

- **Demonstrating the impact of family planning programs on contraceptive use requires more than the simple tracking of contraceptive prevalence over time in a given country.**

Many working in family planning would like to think that "if contraceptive prevalence increases, the program must be successful." However, factors other than the program may have contributed to these increases. Controlled field experiments to demonstrate what would have happened in the absence of the family planning program are not feasible in evaluating ongoing, national level programs. The single largest methodological challenge for evaluating family planning programs in the past decade has been the issue of establishing attribution. (For a full description of the issue and proposed methods of addressing it, see Bertrand, Magnani, and Rutenberg, 1996, Chapter IV). Evaluation methodology is far more advanced for family planning than for other areas of reproductive health, in part because a single, valid outcome indicator measurable through DHS-type studies is available. Yet, definitively establishing cause-and-effect is still relatively rare in the evaluation of family planning programs.

- **The long-term outcome variable for family planning programs is no longer clear-cut.**

Prior to the 1994 Cairo Conference, one of the primary goals of many family planning programs – especially in Asia – was to reduce fertility, an indicator that is reliable and relatively easy to measure. Whereas many governments worldwide continue to track fertility as a desired outcome of family planning programs, a number of programs are repositioning family planning within the larger context of reproductive health as a reproductive right or health intervention. To date, there is no standardized, widely accepted measure of these alternative outcomes. Moreover, these outcomes are highly influenced by other factors in addition to contraceptive use. The HARI index (**Helping Individuals Achieve their Reproductive Intentions**) has been proposed to fill this void (Jain, 1995; Jain, 2001), but is not yet in widespread use. In short, the field has moved away from a single-minded focus on fertility, but an alternative indicator that reflects the health and reproductive rights aspects has yet to surface.

Almost all of the indicators in this section were taken directly from the *Handbook of Indicators for Family Planning Program Evaluation* (Bertrand, Magnani, and Knowles, 1994), suggesting little change in the basic indicators for evaluating outputs and outcome in family planning programs since that time. However, we have reduced the total number of indicators of family planning/fertility from 35 to 14, retaining those which have proven most useful for program evaluation in field settings.

---

[1] Program methods include oral pills, IUDs, injections, condoms, NORPLANT implants, male sterilization, female sterilization, spermicides, and lactational amenorrhea method (LAM). Non-program methods include rhythm, withdrawal, abstinence, and folkloric methods. The lower effectiveness of certain methods is reflected in the way major surveys treat them. The DHS collects data on all the above-cited methods, but future reporting of contraceptive prevalence may exclude abstinence and folkloric methods. The RHS surveys omit LAM, abstinence, or folkloric methods in their calculation of contraceptive prevalence. Note: The distinction between program and non-program methods varies by country. For example, some programs directed at youth promote abstinence, which in this context is a program method.

FAMILY PLANNING PROGRAM EFFORT INDEX

Definition
This indicator is a score (ranging from 0-120) measuring the strength of the national family planning program of a given country on four dimensions (policy and stage setting activities, service and service-related activities, evaluation and record keeping, availability and accessibility of fertility control supplies and services). The score has a potential range of 0-120 points, based on 0-4 points for each of 30 items.

Data Requirements
Responses to a detailed questionnaire from selected key informants (representatives of the Ministry of Health, IPPF affiliate or other NGO; international consultants familiar with that country; and other informed individuals)

Data Source(s)
A questionnaire designed explicitly for this purpose, completed by an average of four to six individuals per country (with extreme cases ranging from 1-12 respondents). The items have remained constant over the four rounds of data collection.

Purpose and Issues
The purpose of the Family Planning Program Effort Index is to assess the strength of family planning programs worldwide and to measure changes over time. It attempts to measure the effort (input) that goes into the family planning program, not the results achieved. These data have been collected in five different cycles over a 27-year period from 1972 to 1999. The questionnaire consists of some 125 different questions that relate to 30 different dimensions of program effort. Researchers convert the responses to these questions to individual scores (ranging from zero to four) for each of the 30 items, using an established set of rules.

The Family Planning Program Effort Index serves several important purposes:

1. It allows for cross-national comparisons of family planning programs at different points over the past 27 years;
2. It traces the evolution of the family planning program in a given country over time;
3. It measures family planning program input, independent of outcomes (such as contraceptive prevalence or fertility). As such, it has been used to analyze the relative importance of family planning program effort versus social and economic factors in the decline of fertility rates worldwide (Lapham and Mauldin, 1984; Mauldin and Ross, 1991; Ross and Mauldin, 1996; Ross and Stover, 2001).

This index is useful primarily to researchers, donor agency representatives, and those interested in understanding family planning programs in the international context. Although it indicates areas of strength and weakness for a national program, it has not been the tool of preference for program managers at the country level in identifying ways to improve programs. Indeed, in contrast to other widely used measures and tools (CYP, contraceptive prevalence, DHS and RHS surveys, situation analysis) which are widely known to family planning program managers throughout the developing world, the Family Planning Program Effort Index is used primarily for research purposes.

Two major criticisms have been leveled against this index. First, some critics argue that a sample of four to six key respondents per cycle of data collection is inadequate to accurately capture the complexity of national family planning programs. Some countries have been scored using data from as few as one or two respon-

dents. Second, many argue that the responses of the key informants are biased by their knowledge of key outcomes (contraceptive use and fertility decline). That is, if contraceptive prevalence is high, the respondents unconsciously give high scores to the availability of methods, assuming that the two go hand in hand. In short, although the index claims to measure inputs, the responses may be biased by a knowledge of outcomes.

Despite these limitations, the Family Planning Program Effort Index represents a valuable source of information to the international reproductive health community. It is the only source of data that purports to measure inputs using a standard set of questions across countries and over time. The results for the Family Planning Program Effort Index have generally coincided with qualitative assessments of "how good a family planning program is" in a given country. For example, the scores indicate that in 1972 a quarter of the world's population lived in countries with very weak or no family planning programs. As of 1999, no country in the world fell in this category. Conversely, the percentage of the world's population living in countries with strong family planning programs increased from 36 percent in 1972 to 68 percent in 1999 (Ross and Stover, 2001). Curiously, some of the strongest family planning pro-grams in the world (e.g., in China and the Republic of Korea) actually dropped in the most recent cycle of data collection, presumably because the established norm for small families and contraceptive use no longer required the aggressive family planning program initiative of previous years.

Most of the indicators in this *Compendium* are included to encourage evaluators and researchers to collect the necessary data to use them. By contrast, the data collection for the Family Planning Program Effort Index is conducted by a small team of researchers based in the United States (which has included Bernard Berelson, Robert Lapham, Parker Mauldin, John Ross, and John Stover over different cycles), using a standardized instrument across countries. Further validation of the FP Program Effort Index is available from Mauldin and Ross, who demonstrated that the "judgmental" scores of key informants were relatively consistent with more direct measures of the different components in a case study that specifically investigated their comparability (Mauldin et al., 1995). The index is useful to the international reproductive health community because it offers data of this nature for secondary use, not as a type of information routinely collected by program managers for the purpose of program improvement.

HELPING INDIVIDUALS ACHIEVE THEIR REPRODUCTIVE INTENTIONS (HARI INDEX)

Definition

The effectiveness of family planning programs from the user perspective of successfully achieving reproductive intentions, rather than the more demographically oriented indicators of contraceptive prevalence or TFR

The index consists of two components: the achievement of an individual's reproductive intentions and the avoidance of severe reproductive health problems associated with an individual's effort to achieve her stated reproductive intentions (Jain, 2001).[2]

Data Requirements

Self-report of the following information by women of reproductive age:

- Reproductive intentions and contraceptive use at two time periods, and the occurrence of pregnancies and births during the interim period;
- Sterilization regret; and
- Experiences with severe reproductive health problems associated with pregnancy and contraception during the interim period.

Data Source(s)

A panel survey with two rounds of data collection among the same respondents (ideal for estimating the achievement of **individuals'** reproductive intentions); cross-sectional surveys (possible, but as yet untested in the literature)

Purpose and Issues

Consistent with the goals set at the 1994 ICPD in Cairo, this indicator broadens the scope of family planning programs to incorporate reproductive health services. Whereas family planning programs have traditionally been evaluated in terms of contraceptive use and fertility reduction, the objective of many current programs is to help individuals achieve their reproductive intentions in a healthful manner.

Jain proposed the HARI Index in 1994, but until his article in 2001, no one had published results of a field test of this indicator. Experience in the use of this indicator is limited to this single panel study in Peru, consisting of two rounds of data collection 29 months apart. The standard DHS questions were expanded slightly to allow for more in-depth information related to the experience of sterilization and potential infertility problems. Additional questions were also added in the second round to retrospectively collect information on reproductive health problems experienced by women during the interim period. Although this indicator is relatively new, we include it in the *Compendium* because it is one of the few indicators that attempts to capture the goals of reproductive choice expressed at the Cairo conference.

One could measure the two components of HARI separately. However, the potential advantage of using a combined indicator is to draw the attention of researchers, service providers, and program managers to both components. Although severe health problems associated with contraception and pregnancy in this application tend to contribute less to the HARI index than do mistimed pregnancies, such problems are extremely important from the individual's perspective. The HARI index has several strengths. First, it measures success and failure of a reproductive health program from the user's perspective of successfully achieving RH intentions, consistent with the Cairo agenda. Second, a program is considered successful if the woman has a wanted and intended pregnancy. Third, it is considered successful if a woman avoids an unwanted or mistimed pregnancy, regardless of her contraceptive use or her visits to facilities to avoid an unintended pregnancy. The index reflects the realization of a woman's reproductive intentions stated by those surveyed. As such, the index overcomes the problem created by clandestine or unreported use of contraceptives identified in some settings. Moreover, the index draws attention to the need to shift

---

[2] The description of this indicator draws heavily on the article by Anrudh Jain (2001).

attention from reducing wanted pregnancies to avoiding severe reproductive health problems when one designs family planning programs.

This current application of the index has several potential limitations. First, considerable overlap exists between the two components of HARI: the achievement of individual reproductive intentions and the avoidance of severe health problems associated with an individual's effort to achieve her stated RH intentions. Thus, the single figure does not fully reflect reasons women did not achieve their reproductive intentions in a healthful manner, although presentation of data by each subcomponent readily addresses this shortcoming.

Second, the HARI Index is best applied in the context of a panel survey (i.e., with the same respondents on two rounds of data). Panel surveys provide a more accurate estimation of the achievement of an individual's reproductive intentions component of the index, but they are rarely conducted in developing countries. The items concerning reproductive intentions and the wantedness of recent pregnancies or births are standard content in demographic surveys. However, questions regarding serious health consequences related to pregnancy and contraception are not standard to the existing RH surveys (i.e., DHS, RHS). Including such questions on these RH surveys would allow evaluators to estimate this indicator from cross-sectional surveys on a large scale.

Third, each item in the HARI index receives equal weight. For example, having an unwanted birth counts the same as experiencing a reproductive health problem or having a mistimed pregnancy. Moreover, a woman who fails to achieve her reproductive health intentions on both components of the score (e.g., having both an unwanted pregnancy and a hospitalization due to abortion-related complications) may have the same score on the HARI as one who fails on only one of the components. Also, the index is not adjusted for changes in reproductive health intentions or inappropriate provisions or inappropriate use of reversible methods. HARI can be adjusted by using differential weights for women with different problems and by counting inappropriate provision or use of contraception as failure, provided such information is collected.

Finally, the current application of HARI cannot assess the relative importance of a larger set of reproductive health problems outlined in the ICPD Programme of Action. We gain little by incorporating well-established indicators of infant, child, and maternal mortality in the HARI index. Breast and cervical cancers are too rare to be captured in a survey of this kind. Progress on preventive actions can be monitored through such process indicators as the percentage of women visiting a clinic who are informed about breast examination, or who have a mammogram or a Pap smear test on schedule. HARI does incorporate a smaller but important set of reproductive health problems associated with pregnancy and the practice of contraception. HARI can incorporate other important reproductive morbidities (e.g., reproductive tract infections and STIs) as progress is made toward identifying them in the absence of lab tests.

Number of Acceptors New to Modern Contraception

Definition
The number of persons who accept for the first time in their lives any (program) method of contraception; to be reported for a defined reference period (e.g., one year)

Data Requirements
Counts of persons accepting any (program) method for the first time in their lives during a one-year period

Data Source(s)
Service statistics; surveys (possible but uncommon)

Purpose and Issues
This indicator measures the ability of the program to attract new clients from an untapped segment of the population to its services. The measure eliminates the problem of counting as "new" those clients who switch from one source to another for reasons of convenience or cost. As an indicator, it may also reflect the success of special communication programs or other interventions (e.g., social marketing projects) aimed at increasing service utilization among those previously missed by the program. However, in this latter case, one must be mindful that some of the new acceptors might have obtained the same or another method from an alternate source (e.g., the unsubsidized pharmacy sector) if the special intervention had not taken place.

"Program method" refers to methods made available through established family planning programs: pill, IUD, the NORPLANT implant, injection, condom, spermicides, diaphragm, tubal ligation, vasectomy, and lactational amenorrhea method (LAM), if used under program supervision. Thus, a young woman who formerly obtained condoms from the pharmacy is not a new acceptor. By contrast, a client who to date has depended on withdrawal is a new acceptor, because withdrawal is not a program method.

The **Number of Acceptors New to Modern Contraception**, defined as first-time use in the life of the individual, reduces the ambiguity associated with the more general term "new acceptor" and avoids a duplication of cases that may result when substitution occurs.

Evaluators can obtain this indicator from survey data as well (e.g., from the "calendar" used in the DHS or other data collection tools for obtaining contraceptive histories retrospectively). Moreover, surveys allow one to include non-program methods. However, surveys are rarely used to produce data on acceptors, and total current use rather than "new use" is likely to be of greater interest to those interpreting the data.

Program personnel (including evaluation staff) can disaggregate service statistics by key variables (age, sex, parity, place of residence, ethnicity, or other factors judged relevant in the country context) to obtain a socio-demographic profile of the client population. This information is useful in tracking changes in the composition of the client population over time and in determining whether programs intended to reach certain subgroups are effectively doing so.

COUPLE-YEARS OF PROTECTION (CYP)

Definition

The estimated protection provided by family planning services during a one-year period, based upon the volume of all contraceptives sold or distributed free of charge to clients during that period

The CYP is calculated by multiplying the quantity of each method distributed to clients by a conversion factor, to yield an estimate of the duration of contraceptive protection provided per unit of that method (Wishik and Chen, 1973; Stover, Bertrand, and Shelton, 2000). The CYPs for each method are then summed over all methods to obtain a total CYP figure.

The EVALUATION Project undertook an extensive review of the literature and empirical data on a number of the variables that form the underlying assumptions for the calculation of CYP. Based on this analysis, the researchers recommended a modified set of conversion factors for CYP (described in Stover, Bertrand, and Shelton, 2000). USAID accepted this set of recommended factors but "harmonized" them with the previous set of conversion factors (i.e., they retained the old factors when they were close to the new ones to minimize disruption to data collection systems worldwide). USAID then issued a slightly modified set of conversion factors, which the USAID system has used since 1997. The CYP conversion factors endorsed by USAID are as follows:

Method	CYP Per Unit
Oral Contraceptives	*15 cycles per CYP*
Condoms	*120 units per CYP*
Vaginal foaming tablets	*120 units per CYP*
DepoProvera (injectable)	*4 doses (ml) per CYP*

Noristerat (injectable)	*6 doses per CYP*
Cu "T" 380–A IUD	*3.5 CYP per IUD inserted*
NORPLANT implant	*3.5 CYP per implant*
Sterilization(male or female)	*9 CYP per procedure (global default value)**
Natural Family Planning:	*2 years per trained, confirmed adopter*
Lactational Amenorrhea Method (LAM)	*4 active users per CYP*

* Note: Because of marked differences in CYP for sterilization by country and by region (based on differences in median age at sterilization), countries should use the median value for their region (assuming their data on age at sterilization conform to those of the region). The regional values for one sterilization are:

- Asia   10 CYP
- Latin America   10 CYP
- Africa   8 CYP
- Near East/North Africa 8 CYP

Programs wishing to use country-specific statistics are referred to the Stover, Bertrand, and Shelton (2000) report for the appropriate CYP.

An illustrative computation of this indicator is provided at the end of the discussion of the indicator.

Data Requirements

Quantities of pills, condoms, and spermicides distributed to clients; numbers of IUDs and NORPLANT implants inserted; number of injections administered; number of sterilization operations performed; number of trained, confirmed clients of NFP; number of LAM clients during the reference period

Data Source(s)
Service statistics or logistics management information system

Purpose and Issues
CYP measures the volume of program activity. Program managers and donor agencies use it to monitor progress in the delivery of contraceptive services at the program and project levels. Because USAID and IPPF generally require the organizations they support to report CYP, this measure is currently one of the most widely used indicators of output in international family planning programs.

This indicator has several advantages:

- It can be calculated from data routinely collected through programs or projects, and thus minimizes the data collection burden;
- These data can be obtained from all the different service delivery mechanisms (clinics, CBD, social/commercial marketing); and
- The CYP calculation is relatively simple to do.

The principal disadvantages of the indicator are that:

- It is not intuitively easy to understand by those outside the field;
- One cannot ascertain the number of individuals represented by CYP. For example, if a program administers 10,000 injections of DepoProvera, this amount is equivalent to 2,500 CYP. Theoretically, this figure represents 2,500 women protected for 12 months each; however, in fact it may refer to 5,000 women covered for 6 months each or 10,000 women covered for 3 months each; and
- The validity of the assumptions underlying the choice of conversion factors is open to debate.

Regarding the calculation of CYP for long-term methods, most programs "credit" the entire amount to the calendar year in which the client accepted the method. For example, if a family planning program performed 100 voluntary surgical contraception (VSC) procedures in a given year, it would credit all 1000 CYP (100 procedures x 9 years/each) to that calendar year, even though the protection from those procedures would in fact be realized over that and the next nine years. An alternative approach is to "annualize" this projection, allocating it over a nine-year period. The same principle applies to IUDs and the NORPLANT implant. Although the first approach (of crediting the full amount of CYP in the calendar year of acceptance) has been harshly criticized, it represents current practice in most programs that report CYP, probably because it is easier to apply.

Ideally, CYP should be based on the volume of contraceptives delivered to clients who will presumably use them, not on those delivered to facilities where they may remain unused in cartons or on shelves. However, in some projects such as social marketing, it may be impossible to monitor the exact numbers reaching the hands of clients. Rather, the only means of calculating CYP is to base it on the volume of contraceptives delivered to the retailers in question. Given that retailers are unlikely to stock products that move slowly, it is probable that (after an initial shipment) most contraceptives sold to retailers will make their way into consumers' hands. However, in those instances where the calculation of CYP is based on the volume of products delivered to retailers, not directly to the clients or customers themselves, those preparing the CYP report should clarify this detail to the users of the information.

Illustrative Computation
CYP, based upon conversion factors given in text

Method	Quantity	CYP
Oral contraceptives	5,022	334.8
IUDs	87	304.5
Condoms	62,810	523.4
Vaginal tablets	3,900	32.5
Tubal ligations	13	117.0
DepoProvera	1,277	319.3
TOTAL		1631.5

METHOD MIX

Definition
The percent distribution of contraceptive users (or alternatively, of acceptors) by method

Data Requirements
Number of users (or acceptors) by method

Data Source(s)
Service statistics (program-based) or DHS-type surveys (population-based)

Purpose and Issues
The method mix provides a profile of the relative level of use of different contraceptive methods. A broad method mix suggests that the population has access to a range of different contraceptive methods. Conversely, method mix can signal: (1) provider bias in the system, if one method is strongly favored to the exclusion of others; (2) user preferences; or (3) both.

Method mix often changes in response to the introduction of a new method in-country (e.g., the injection), to non-availability of methods due to stockout, to increased need for a method that also protects against STIs (i.e., condoms), and to user preferences. Data on method mix can signal these changes, but do not provide insight into the reasons for the change. Evaluators can use qualitative methods to better understand the clients' motivations for switching methods.

Because of the problems of monitoring the number of current users based on service statistics, method mix is generally based on acceptors, not on current users, when measured at the program level. The two yield different distributions, since user data reflects the accumulation of long-acting methods from previous years.

Similarly, one expects some discrepancy on method mix calculated from program statistics versus surveys, even in programs with reliable data. (The reason is that program-based statistics reflect activity in the calendar year under study, whereas the survey results include continuing users of long-acting methods who adopted them in previous years and have not needed or chosen to return to the clinic in the calendar year under study). In addition, survey data may include folk methods, non-program methods (e.g., withdrawal), and program methods also available from non-program sources (e.g., pills from pharmacies).

In the case of method mix, the question is not which source of data is better: program- versus population-based. Both are used in forecasting the future contraceptive needs of a country. Many evaluators consider survey data more reliable for assessing preferences for specific methods, because they include clients from both the public and private sector, in addition to those using a non-program method such as withdrawal. However, one must be mindful that in survey data (e.g., the DHS or RHS) the coefficient of variation may be large, and thus affect the stability of the estimate, especially where the percentage using a specific method is very low. Finally, survey data and service statistics sometimes differ, a situation that can arise from inflated service statistics, wastage in the system, or the sale of products outside the intended area for the program (e.g., across borders).

Despite considerable discussion of method mix, there has been relatively little published research on what constitutes a desirable method mix, exceptions being Hutchings, Perkin, and Saunders (1987) and Potter (1999). Practitioners generally feel that a program should respond to the changing needs of the population at different stages in the reproductive life cycle, and offer reversible methods for those who desire to space pregnancies and permanent methods for those who have completed their desired family size. Thus, programs offering no permanent methods or overemphasizing permanent methods are subject to criticism. Yet within the category of reversible methods, the distribution of acceptors by type of contraceptive will vary by availability of specific methods, costs, local preferences, and other factors, and thus make it difficult to generalize regarding desirable method mix.

Gender Implications of this Indicator

Contraceptive method mix can be one indication of gender balance in contraceptive responsibility within a country or program. Globally, vasectomy, though safer and less costly, is much less widely available and used than female sterilization is. Nearly a third of all contraceptors rely on female sterilization, while only seven percent rely on vasectomy. In India, the ratio of 35 female sterilizations to 1 vasectomy suggests that the program is heavily biased towards female responsibility for contraception. In many parts of sub-Saharan Africa, vasectomy remains virtually unknown. The condom, rhythm, and withdrawal also require male participation or responsibility. Encouraging greater gender equity in contraceptive practice is one goal of the efforts to involve men as partners in family planning and reproductive health.

CONTRACEPTIVE PREVALENCE RATE (CPR)

Definition
The percent of women of reproductive age who are using (or whose partner is using) a contraceptive method at a particular point in time, almost always reported for women married or in sexual union

Generally, the measure includes all contraceptive methods (modern and traditional), but it may include modern methods only.

The indicator is calculated as follows:

$$\frac{\text{\# of women ages of reproductive age (married or in union) using a contraceptive method}}{\text{Total \# of women of reproductive age (married or in union)}} \times 100$$

Illustrative Computation
For example, the DHS for Peru (2000) yielded the following data on CPR, among women 15-45 years of age:

All women	**Women currently married or in union:**
CPR= 12,240/27,843	CPR = 10,764/15,628
=0.440 x 109	= 0.689 x 100
=44.0	= 68

Source of data: Peru Demographic and Health Survey, 2000

Data Requirements
The total number of women of reproductive age, by marital status; and of these, the number that are currently using a contraceptive method

Data Source(s)
Population-based surveys

Purpose and Issues
The CPR provides a measure of population coverage of contraceptive use, taking into account all sources of supply and all contraceptive methods; it is the most widely reported measure of outcome for family planning programs at the population level.

Technically speaking, CPR is a ratio, not a rate. Prevalence is measured by a ratio and incidence by a rate. For a given year, contraceptive prevalence measures the percentage of women of childbearing age in union who use a form of contraception. To obtain a true contraceptive use rate, the denominator should reflect the population at risk (of pregnancy), i.e., sexually active women who are not infecund, pregnant, or amenhorreic. The numerator should reflect the number of contraceptive users from that population. Note: We include this point for informational purposes only. The international population community uses the term "contraceptive prevalence rate" as defined above; thus, this *Compendium* endorses this practice to assure consistency.

The convention in reporting contraceptive prevalence is to base this calculation on women married or in sexual union (even though most DHS-type surveys ask questions of contraceptive use to women of reproductive age, regardless of their marital status). In countries with relatively little sexual activity outside marriage for women, basing prevalence estimates on women in sexual union captures the population at risk of pregnancy. However, in countries with the widespread practice of sexual activity outside of marriage or stable sexual unions, a prevalence estimate based on women in union only would ignore a considerable percentage of current users. Thus, researchers and program evaluators generally report percentage of sexually active unmarried women using contraception, if appropriate, in addition to contraceptive prevalence, because method mix is very different for those married versus unmarried (in/not in a stable union).

Whereas evaluators may theoretically derive the CPR from service statistics on numbers of current users and estimates of the population at risk, current practice is to rely upon population-based sample surveys in order to minimize the problems associated with maintaining a running count of current users and with obtaining accurate population estimates. (The problems include incomplete data, double-counting of users who enter the service delivery system at more than one point, purposeful inflation of service statistics, and poor quality of data due to other activities competing for the attention of those recording the information, to name the primary ones.)

The DHS and RHS are currently the main sources for obtaining national level estimates of prevalence. As mentioned in Part I, "DHS" is used in this *Compendium* to mean "DHS-type surveys:" the actual DHS, the RHS surveys conducted with technical support from CDC, and other large-scale national surveys conducted by the countries themselves under other auspices (e.g., in Algeria, Bangladesh, China, Costa Rica, Cuba, Hong Kong, India, Singapore, South Africa, South Korea, Taiwan, Turkey, and Vietnam). Evaluators may also use smaller scale and/or more focused surveys to estimate the CPR as long as they use probability sampling methods, the essential ingredient for obtaining scientifically-sound estimates. Evaluators may also obtain CPR by adding relevant questions to surveys on other topics (e.g., health program prevalence or coverage surveys), assuming appropriate sampling methods and sample sizes.

SOURCE OF SUPPLY (BY METHOD)

Definition
The percent distribution of the types of service-delivery points cited by users as the source of their current contraceptive method (if more than one source, then the most recent one)

Data Requirements
Number of respondents currently using contraception, the type of method used, and the source of supply of their method (most recently)

Data Source(s)
Population-based surveys

Purpose and Issues
This indicator is useful to family planning program officials because it shows where contraceptive users obtain their supplies and thus allows programs to evaluate their effectiveness and to forecast procurement needs. It is particularly appropriate to countries trying to shift the burden for family planning services from the public to the private sector. For example, the DHS-type surveys yield information on the percentage of modern-method prevalence accounted for by the private sector.

In most countries, the source of supply will vary substantially by type of method. VSCs, IUDs, and the NORPLANT implants require a clinic-based facility (including mobile clinics). Pills are available through clinics in addition to commercial and CBD outlets. DepoProvera, once a clinic-based method, has been introduced into CBD programs and is available in pharmacies in some countries. Condoms and spermicides can be dispensed from any type of facility. Thus, data on source of supply are particularly useful when classified by method.

"Source of supply" yields two types of information: type of facility and type of sector (public/private). Type of facility generally includes hospital, health center, family planning clinic, mobile clinic, pharmacy, field worker, private doctor, and shop, among others. Sector distinguishes between governmental programs and those in the private sector (including the local family planning association, commercial retailers, private physicians, and other private providers). Ideally, data on source of supply should yield the percentage of contraceptive use attributable to the government program, the private family planning association, the private sector (pharmacies, private doctors), and other relevant sources.

However, the distinction between public and private is often difficult to make, especially in countries with multiple sources of contraception. The respondent may incorrectly identify a given clinic as a government clinic, when in fact it is private (or she simply may not know if it is public or private). A private physician may in fact be participating in a subsidized program to offer low-cost services to specific groups. In response to this problem, the DHS questionnaire now provides a line for entering the actual name of the facility. Subsequent to the interview, a member of the research team codes the place mentioned according to the correct classification, based on master lists of SDPs. To classify those SDPs not on the list, researchers can later contact key informants from the area.

CONTINUATION RATES

Definition

The cumulative probability that acceptors of a contraceptive method will still be using any contraceptive method offered by the program after a specified period of time (e.g., one year)

This is also known as the "all-method" continuation rate.

When using cross-sectional population data, evaluators calculate the continuation rate for each unit-interval of use (e.g., first, second, third month of use, and so forth) as the complement of the ratio of acceptors who discontinue use of a program method of contraception at that duration to the number of women still using at the beginning of the month (i.e., 1 minus the discontinuation rate). Evaluators then cumulate these continuation rates to obtain the probability that acceptors of a contraceptive method will still be using any program method after the specified period of time.

The indicator (CR_x) is calculated as:

$$CR_x = \frac{x\ (1-q_x)}{\Pi}$$

Where:

$x = 1$

$q_x = T_x/N_x$ = conditional probability of discontinuing use during a given interval (e.g., one month, one quarter);

T_x = the number of women discontinuing use during the interval; and

N_x = number of women using at the beginning of the interval.

Note: Π signifies that $(1-q_x)$ is multiplied over all intervals from 1 to x.

Illustrative Computation

Continuation rates for reversible methods for durations from 1-12 months, Bangladesh, 1992-97

x	T_x	N_x	q_x	CR_x
1	291.3	4422.5	.0659	93.4
2	160.1	4025.9	.0398	89.7
3	205.9	3803.1	.0541	84.8
4	89.9	3502.6	.0257	82.7
5	61.2	3322.9	.0184	81.1
6	107.5	3173.7	.0339	78.4
7	68.7	2998.3	.0229	76.6
8	57.2	2868.5	.0199	75.1
9	74.6	2757.4	.0271	73.0
10	59.8	2579.8	.0232	71.3
11	55.8	2458.0	.0227	69.7
12	109.5	2350.8	.0466	66.5

Source of data: Bangladesh Demographic and Health Survey, 1996/97

Data Requirements

Information on contraceptive initiation, duration of use (including method switching), and discontinuation during a given reference period (e.g., the 3-5 years prior to a survey). Based on this information, one can calculate the percentage who have continuously used for a specific duration (e.g., 12 months, 18 months), as well as the median duration of use.

Data Source(s)

Population-based: surveys with retrospective contraceptive use histories or calendars

Program-based: client records accompanied by a follow-up study of program dropouts. This source is rarely used.

Purpose and Issues

Contraceptive continuation rates provide a useful summary measure of the overall effectiveness of program

services in enabling clients to sustain contraceptive use even though they may switch from one method to another. However, the calculation of continuation rates from surveys requires knowledge of life table (survival) analysis by those subregions of the country (much less individual facilities), making this indicator more useful at the national than regional or local level.

Although evaluators can calculate continuation rates from either facility-based or population-based data, facility-based data have a number of limitations; thus, researchers tend to use large-scale surveys to provide more valid measurements of continuation among the intended population (e.g., Blanc, Curtis, and Croft, 1999).

Obtaining continuation rates at the program level is theoretically possible if evaluators use follow-up studies of new acceptors at a specified period of time after adoption of the method (e.g., 12 months). However, this technique is rarely used (except in clinical trials), given the difficulty and expense of locating these acceptors a year later.

The preferred source of data is the "calendar," a data collection format used in cross-sectional surveys such as the DHS. However, such surveys have limitations of their own. They (a) depend upon the accuracy of respondent recall, (b) do not allow linking of respondents to specific SDPs, and (c) may not capture the full contraceptive history (e.g., when five-year calendar is used).

It is important to note the distinction between discontinuation and failure of a contraceptive method. Discontinuation of contraception may occur because the individual chooses to stop using a selected method or because accidental pregnancy intervenes. As such, method failure is a subset of discontinuation. Method failure necessarily results in discontinuation. However, not all discontinuation is attributable to method failure.

Indicator

Unmet Need for Family Planning

Definition

The number or percent of women currently married or in union who are fecund and who desire to either terminate or postpone childbearing, but who are not currently using a contraceptive method

The total number of women with an unmet need for family planning consists of two groups of women: (a) those with an unmet need for limiting, and (b) those with an unmet need for spacing.

Women with an unmet need for limiting are those who desire no additional children and who do not currently use a contraceptive method. Women with an unmet need for spacing are those who desire to postpone their next birth by a specified length of time (for example, for at least two years from the date of a survey) and who do not currently use a contraceptive method.

The indicator is calculated as follows:

$$U_L + U_S = U$$

Where:

U = the number or percent of women with unmet need for family planning;

U_L = the number or percent of women with an unmet need for limiting; and

U_S = the number or percent of women with an unmet need for spacing.

Note: The actual calculation of unmet need is fairly complex, as described in detail in Appendix H.

Illustrative Computation

Estimate of unmet need for family planning, Peru, 2000 (expressed as the percentage of women currently married or in union).

$$U = U_L + U_S$$
$$= 6.7 + 3.5$$
$$= 10.2$$

Source of data: Peru Demographic and Health Survey, 2000

Data Requirements

Responses to survey questions on:

- Desire for additional children and, if so, the desired length of birth interval;
- Current contraceptive use status;
- Current fecundity, pregnancy, and amenorrhea status for women not currently using a contraceptive method;
- The planning status (with respect to number and/or timing) of the current/last pregnancy for women currently pregnant or amenorrheic; and
- Use (or not) of a contraceptive method at the time of the current/last pregnancy.

Note: The use of the information in the final two items in the computation of the indicator is explained below.

Data Source(s)

Population-based survey

Purposes and Issues

This indicator provides information on the size of an extremely important population sub-group for family planning program management: women at risk of pregnancy with an apparent need for family planning services based upon their expressed desire to limit or space future births, but who do not use contraception. Such women have an "unmet demand" or "unmet need" for family planning and are the logical primary audience of program efforts.

223

The indicator may also be interpreted as the number of additional clients who would be using contraception (over and above the number of current users) if all women at risk of pregnancy and desiring to either terminate or postpone childbearing were to adopt contraception.

The indicator follows from the breakdown of total demand for family planning services into two components: "met demand" and "unmet demand" (or "unmet need"). Met demand consists of women with demand for family planning who are using a contraceptive method to achieve their reproductive goals; unmet need, or unmet demand, consists of women with an apparent demand for family planning who are not using contraception.

Following the procedure proposed by Westoff and Ochoa (1991), women are considered to be at risk of pregnancy in the present indicator if they are:

- Of reproductive age and currently married or in union;
- Fecund;
- Not using a contraceptive method; and
- Not currently pregnant or amenorrheic.

However, the following categories of women are not considered to have an unmet need for family planning, and thus, when computing the indicator, evaluators should exclude:

- Currently pregnant or amenorrheic women who were using contraception at the time they became pregnant with the current/last birth (these women are viewed as not in need because prior need was met through contraceptive use, although they do appear to need a more effective method);
- Currently pregnant or amenorrheic women whose pregnancy was reported as intentional; and
- Fecund women who want their next child within the next two years.

Bongaarts (1991) has proposed two modifications to the measurement procedure described in Appendix H:

(1) an adjustment to account for the fact that the satisfaction of need for spacing through contraceptive use will reduce the need for limiting to the extent that it postpones the date at which women reach their desired family size; and (2) an adjustment for a perceived over-estimate of the need for spacing in the procedure described above. Bongaarts proposes that evaluators use the estimates produced by the procedure described above as a starting point and introduce these two adjustments to compensate for the perceived problems. A comparison of the estimates from the two methods suggests that the Westoff procedure tends to produce estimates of the level of unmet need that are higher than those of Bongaarts by, on average, about five percent (Bongaarts, 1991; Westoff and Ochoa, 1991). The reader is referred to these references for further details on the two methods of computing the indicator.

Another refinement consists of reporting unmet need for married women, unmarried women, and all women of reproductive age. This refinement is particularly relevant for countries in which a significant share of childbearing occurs outside of recognized marriages/unions.

Some researchers have argued that the definition of "unmet need" should be broadened to include women using: (1) traditional contraceptive methods (on the grounds of high failure rates for such methods); (2) a theoretically effective method incorrectly or sporadically; and (3) a method that is unsafe or unsuitable for them (Foreit, 1992; Dixon-Mueller and Germain, 1992). The RHS has modified the calculation of unmet need to include traditional contraceptive methods in countries where they are in widespread use (e.g., Eastern Europe, Turkey, Mauritius). The adoption of these alternative definitions would raise significantly the estimated numbers of women with unmet need for family planning in many developing country settings.

A related indicator, the satisfaction of demand for family planning services, consists of the percentage of total demand for family planning at any time that is being satisfied by current contraceptive use. Thus:

Satisfaction of demand for FP = 1 – unmet need

Gender Implications of this Indicator

A gender-sensitive approach to unmet need would examine which factors lead to unmet need, distinguish between the unmet need of women and men, and include gender-sensitive service-delivery strategies.

1. Factors that lead to unmet need:
   - Do women and men have different access to the knowledge and household resources that would enable them to use family planning effectively?
   - Do women and men have different levels of decision-making autonomy and freedom of movement that would enable them to use family planning effectively?
   - Do women and men have the communication skills to discuss their fertility and contraceptive preferences with their partners?
   - Is family planning use a factor in gender-based violence, actual or feared?

2. Unmet need of women and men:
   - To what extent are fertility preferences shared between women and men?
   - Are cultural norms regarding extramarital sexual relations different for women and men, and the expectations of bearing children with different sexual partners?
   - In societies with polygamous unions, how do women and men view childbearing?
   - Is son preference a dominant issue in different fertility preferences between women and men?

3. Service-delivery issues:
   - Are providers trained to recognize gender-based obstacles to effective use of family planning (e.g., women clients may find it difficult to ask questions)?
   - Are providers trained to screen for domestic violence?
   - Do providers' own gender-based cultural norms and biases contribute to unmet need, (e.g., unmarried women or widows should not be having sex but it is OK for young men and widowers)?
   - Does the service-delivery system include strategies to mitigate gender-based financial or access constraints? Are services available at times and places convenient to female and male clients?

DESIRE FOR ADDITIONAL CHILDREN

Definition
The number or percent of women (or men) of reproductive age who want to have a (another) child or, conversely, who do not want to have additional children

Data Requirements
Numbers or percent of respondents reporting that additional children are/are not desired

Data Source(s)
Population-based surveys or facility-based data

Purposes and Issues
This indicator is widely used in surveys to identify both: (1) women (or men) with a demand for additional children and (2) those who do not desire additional children and thus have an apparent need/demand for fertility limitation. In the DHS, non-pregnant women married or in union are asked, "Would you like to have a (another) child or would you prefer not to have any (more) children?" Women who are pregnant (or uncertain of their status) at the time of the survey are asked, "After the child you are expecting, would you like to have another child or would you prefer not to have any more children?" On the basis of responses to these questions, evaluators may divide respondents into two categories: those desiring additional children and those desiring to terminate childbearing, with women in the latter category considered as having a "demand for family planning."

Evaluators may also combine responses to this type of question with information on current fecundity and contraceptive use to assess the level of unmet need for family planning (see Appendix H). Comparable information may sometimes be available from service statistics of clinic-based family planning programs. Questions similar to those included in the DHS are often asked of (at minimum) new clients in order to determine the appropriateness of different contraceptive methods in relation to reproductive intentions: that is, methods appropriate for limiting versus spacing.

Despite earlier concerns as to the validity of survey questions of this type in predicting actual fertility behavior, recent studies have provided rather convincing evidence of strong aggregate-level associations between expressed desires for additional children on the one hand and patterns of current contraceptive use and current and future fertility on the other (Bongaarts, 1990; Westoff, 1991). The indicator is currently viewed as relatively unbiased, because respondents have no obvious reasons to over- or under-report preferences to continue childbearing.

Wanted Total Fertility Rate (WTFR)

Definition

The number of children who would be born per woman (or per 1,000 women) if she/they were to pass through the reproductive years bearing children according to a current schedule of age-specific fertility rates if only "desired" or "wanted" births occurred

For this indicator, "wanted" births are defined taking into account both desired family size and the number of surviving children. All births during a specified reference period (usually the two to five years prior to a survey) that do not exceed the respondent's stated "desired number of surviving children" are classified as wanted. Births raising the number of surviving children above the desired family size are considered unwanted.

The indicator is calculated as follows:

$$\text{WTFR} = 5 \sum_a \left(\frac{\text{WB}_a}{\text{E}_a} \right)$$

Where:

WB_a = the number of births to women in age group *a* in a given year or reference period that are "wanted;" and

E_a = the number of person-years of exposure in age group *a* during the reference period.

Data Requirements

Responses to survey questions on:

- Numbers and dates of births during a recent period (typically the two to five years prior to a survey);
- Desired number of children or family size; and
- Number of children ever born and number surviving.

Data Source(s)

Population-based surveys

Purposes and Issues

The WTFR is a measure of "wanted" fertility, a hypothetical measure of what the total fertility rate (TFR) would be given age-specific fertility rates for a recent past period under the condition that all women's fertility preferences were perfectly realized; that is, if only "wanted" births occurred. The measure represents an attempt to avoid the suspected bias in the wanted status of recent births indicator by defining wanted or desired status on the basis of the consistency (or lack thereof) between the reported desired family size and the number of surviving children, instead of on the basis of retrospective reports of fertility intentions at the time of becoming pregnant.

Evaluators calculate the indicator as the sum of age-specific fertility rates, or the total fertility rate, after they delete births occurring during a specified reference period that raise the number of surviving children of sample respondents above their stated desired number of children.

In the DHS, numbers of births during the specified reference period are derived from the birth history portion of the survey interview, the numbers of surviving children are derived from questions on lifetime fertility and survival status, and the information on desired family size are derived from survey questions.

The above definition of the WTFR is based upon the work of Lightbourne (1985, 1987) and Westoff (1991) (who labels the measure the "desired total fertility rate" or DTFR). Bongaarts (1990) proposed a modified definition of the WTFR in which wanted births are defined on the basis of whether survey respondents desired additional births at the time of a survey instead of on the basis of the comparison of the desired number of children and the number of surviving children. Under this definition, births within a specified reference period are

classified as wanted if the respondent reported wanting additional children at the time of a survey.

The argument for the alternative definition is that it is based upon responses to questions on preferences for additional children, an indicator of demand thought to be less affected by reporting biases than the desired family size indicators (Bongaarts, 1990). Comparison of estimates of the two versions of the WTFR for 48 DHS countries indicates that the two measures are reasonably close for most countries, with an average difference between the measures of about 9 percent – 4.09 versus 3.76 (Bongaarts, 1990). On the basis of available evidence, either version of the WTFR is preferable to using the wanted status of previous births in defining wanted fertility.

The comparison of the WTFR with the TFR indicates the extent to which observed fertility exceeds desired or wanted fertility. This type of comparison provides program managers and policy-makers with some insight into the potential short- to medium-term demand for family planning services and the potential for fertility decline in the future (Westoff, 1991). In the case of Burkina Faso, for example, the comparison of the TFR (6.4) with the WTFR (5.7) suggests that a considerable share of current fertility is unwanted and that sufficient latent demand exists in this population; thus an increase in contraceptive prevalence and a decline in fertility might be reasonably expected, if family planning services are available to the population (Institut National de la Statistique et de la Demographie and Macro International Inc., 2000).

AGE SPECIFIC FERTILITY RATES (ASFR)

Definition
The number of births occurring during a given year or reference period per 1,000 women of reproductive age classified in single-or five-year age groups

The ASFR is calculated as:

$$ASFR_a = \frac{B_a}{E_a} \times 1000$$

Where:

B_a = number of births to women in age group *a* in a given year or reference period; and

E_a = number of person-years of exposure in age group *a* during the specified reference period.

Illustrative Computation
Estimates of average annual ASFRs for all women 15-49, Egypt, 1997-2000

Age Group	Births B_a	Person-Years of Exposure E_a	Rate/ Woman	Rate/1,000 Person-Years
15-19	764	14893.2	.051	51
20-24	2304	11747.2	.196	196
25-29	1994	9602.3	.208	208
30-34	1295	8805.5	.147	147
35-39	564	7549.5	.075	75
40-44	161	6643.2	.024	24
45-49	19	4498.8	.004	4

Source of data: Egypt Demographic and Health Survey, 2000

Data Requirements
The number of births in a given year or reference period classified by age of mother and the number of women of reproductive age (i.e., 15-44 or 15-49 years), in 1-or 5-year age groups

Data Source(s)
Vital statistics (numerator only), population censuses or population-based surveys

Purposes and Issues
The ASFR has two primary uses: (1) as a measure of the age pattern of fertility, that is of the relative frequency of childbearing among women of different ages within the reproductive years, and (2) as an intermediate computation in the derivation of the **Total Fertility Rate (TFR)**, discussed next in this section.

As indicated above, evaluators may derive ASFRs from several sources. When evaluators estimate ASFR from vital statistics, they use population projections or estimates of the number of women in each age group 15-49 for the denominator in the rate. When using population censuses or surveys, evaluators obtain both the numerator and denominator of the rate from the census or survey. Estimates from censuses are derived from questions on births during a specified period preceding the census (usually 12 months), while survey estimates may be derived either from questions on births within a specified prior period or from partial or complete birth histories.

A simpler, although less precise, procedure for computing the denominator of the rate is to take the average of the number of women in each age group during the reference period covered by the measure (i.e., the average of the numbers of women in each age group at the beginning and end of the reference period).

Reference periods of more than one year are frequently used to compute ASFRs from survey data, the rationale being to decrease sampling variability associated with relatively small numbers of annual births occurring to women in single or five-year age groups and the distorting effects of reference period reporting errors. Various analyses of DHS fertility data, for example, alternately use the three- or five-year period prior to the survey in calculating ASFRs (Arnold and Blanc, 1989; Lutz, 1990). When multiple years are used for computational purposes, average annual rates are normally presented.

Unlike the crude birth rate, the ASFR is unaffected by differences or changes in population age composition, and thus is more useful in comparing different populations or sub-groups and in measuring changes over time. The ASFR is, however, affected by differences or changes in the number or percent of women exposed to the risk of pregnancy. Thus, changes in ASFRs may provide misleading information regarding the impact of family planning programs on fertility when other factors affecting risk of pregnancy are changing (for example, for the 15-19 and 20-24 age groups when age at marriage is rising quickly).

To address this problem, evaluators may calculate ASFRs only for women who were continuously married or in union during the reference period of the measure. The resulting measure is known as the Marital Age-Specific Fertility Rate (MASFR). However, to calculate this measure, evaluators require data on duration of marriage or marriage histories. In actual practice, MASFRs are more often approximated by calculating ASFRs for women married or in union at the time of a survey, although evaluators should recognize that this figure only approximates the MASFR because women who are married or in union at the time of a given survey may not have been continuously married or in union over the entire reference period of the measure (e.g., for the three to five years prior to the survey).

ASFRs are sometimes presented for different groups of women; for example, ASFRs are for women currently married or in union and for all women of reproductive age in DHS country reports. In societies where fertility is largely confined to marriage, ASFRs for women currently married or in union will provide more or less complete coverage of recent fertility. Where a large share of fertility occurs outside of recognized unions, however, the restriction of the ASFR to currently married women will result in an under estimate of the level of current fertility.

The ASFR is of particular interest in countries, cities, or districts with adolescent RH interventions designed to reduce unintended pregnancy. Although the ASFR is rarely used as an outcome measure in evaluating such programs (due to the human, financial, and logistic resources needed to collect the data), it is a variable that program administrators and policy makers track over time as a macro-level indicator of program effectiveness combined with non-program influences.

Total Fertility Rate (TFR)

Definition
The number of children who would be born per woman (or per 1,000 women) if she/they were to pass through the childbearing years bearing children according to a current schedule of age-specific fertility rates

The TFR is calculated as:

$$\text{TFR} = \Sigma\ \text{ASFR}_a \text{(for single year age groups)}$$

or

$$\text{TFR} = 5\ \Sigma\ \text{ASFR}_a \text{(for 5-year age groups)}$$

Where:

ASFR_a = age-specific fertility rate for women in age group a (expressed as a rate per woman).

Illustrative Computation
Estimate of the average annual TFR for all women aged 15-49, Egypt, 1997-2000.

TFR= 5 (.051 + .196 + .208 + .147 + .075 + .024 +.004)
　　= 3.53
Where: the figures in parentheses are age-specific rates for the 15-19, 20-24, ... , 45-49 age categories, respectively.

Source of data: Egypt Demographic and Health Survey, 2000.

Data Requirements
A current schedule of age-specific fertility rates (ASFRs), for one- or five-year age groups

Data Source(s)
Vital statistics (numerator only), population censuses or population-based surveys

Purposes and Issues
The TFR is the most widely used fertility measure in program impact evaluations for two main reasons: (1) it is unaffected by differences or changes in age-sex composition, and (2) it provides an easily understandable measure of hypothetical completed fertility.

Although derived from the ASFR, a period fertility rate, the TFR is a measure of the anticipated level of completed fertility per woman (or per 1,000 women) if she/they were to pass through the reproductive years bearing children according to the current schedule of ASFRs. We emphasize that the TFR is only a hypothetical measure of completed fertility, and thus women of reproductive age at any given point in time could have completed family sizes considerably different from that implied by a current TFR, should age-specific fertility rates rise or fall in the future.

Because the TFR is derived from a schedule of ASFRs, the comments and caveats regarding the ASFR also apply to the TFR (i.e., method of computation from different sources of data, effects of changing exposure to pregnancy, and implications of computation for currently married versus all women of reproductive age). As was also the case for the ASFR, the TFR may be computed for women who were continuously married or in union during the reference period of the measure in order to decrease the potentially confounding effects of differences in exposure to the risk of pregnancy (to the extent that differences are associated with marital status). This measure is known as the Total Marital Fertility Rate (TMFR).

Note also that whereas the standard age range for the TFR is ages 15-49, TFRs for other age ranges (e.g., 15-34) are sometimes used for analytic purposes, for example, in order to decrease the influences of truncation when examining cohort trends from birth history data.

Unwanted Total Fertility Rate (UTFR)

Definition

The number of unwanted children who would be born per woman (or per 1,000 women) if she/they were to pass through the reproductive years bearing children according to current rates of unwanted fertility

One can derive the unwanted total fertility rate by subtracting the "wanted" component from the total fertility rate.

The indicator is calculated in one of two ways:

(1)

$$UTFR = 5 \frac{(B_{a,u})}{(E_a)}$$

or (2)

$$UTFR = TFR - WTFR$$

Where:

$B_{a,u}$ = the number of births to women in age group a during a given year or reference period that are unwanted;

E_a = the number of person-years lived by women in age group a during the reference period;

TFR = the total fertility rate for a given year or reference period; and

$WTFR$ = the wanted total fertility rate.

Illustrative Computation

Estimate of the UTFR for all women aged 15-49, Egypt, 1997-2000.

$UTFR = TFR - WTFR$
$= 3.53 - 2.85$
$= 0.68$

Source of data: Egypt Demographic and Health Survey, 2000.

Data Requirements

Responses to survey questions, by age of woman, on:

- Numbers and dates of births during a recent period (typically the two to five years prior to a survey);
- Desired number of children or family size;
- Number of surviving children; and
- Desire for additional children.

Data Source(s)

Population-based surveys

Purposes and Issues

The UTFR provides a hypothetical measure of the average number of "unwanted" births a woman or cohort of women would have during her/their reproductive career(s) if they were to follow current schedules of unwanted fertility. As the wanted status of births is based upon reproductive preferences or demand for children, the indicator provides a conceptually direct measure of family planning program impact in enabling women and couples to achieve their reproductive goals (i.e., to avoid unwanted pregnancies).

In the illustrative computation for Egypt, for example, the estimate of .68 indicates that women in Egypt would have on average .68 unwanted births over the course of their reproductive years if current levels of age-specific fertility and demand for children were to prevail throughout the reproductive years of women of reproductive age at the time of the 2000 DHS.

Part III.C
STI/HIV/AIDS

- AIDS Program Effort Index (API)
- Condoms available for distribution nationwide
- Percent of population with accepting attitudes towards those living with HIV
- Percent of population who know HIV prevention methods
- Percent of population with no incorrect beliefs about AIDS
- Percent of population who know methods of preventing mother-to-child transmission of HIV
- Percent of population requesting an HIV test, receiving a test, and receiving test results
- Voluntary counseling and testing centers with minimum conditions to provide quality services
- Percent of pregnant women counseled and tested for HIV
- Antenatal clinics offering and referring for VCT
- Percent of population who had higher risk sex in the last year
- Percent using condoms at last higher risk sex
- Percent of men having commercial sex in the last year
- Percent of young people having multiple partners in last year
- Percent of young people using a condom at last higher risk sex
- Percent of injecting drug users never sharing equipment in the last month
- Percent of transfused blood units screened
- Percent of STI patients appropriately diagnosed and treated
- Percent of STI patients receiving advice on condom use and partner notification and referral to HIV testing services
- Percent of health facilities providing STI services with adequate drug supply
- Number/percent of health facilities with the capacity to deliver appropriate care to HIV-infected patients
- HIV prevalence among pregnant women 15-24 years old
- HIV prevalence in sub-populations with high-risk behavior
- Percent of children under 15 who are orphans

HIV/AIDS has become a major focus for public health specialists and international donor agencies, as this devastating pandemic continues to spread. According to UNAIDS, at the end of 2001, an estimated 40 million people globally were infected with HIV; of these, 70 percent reside in Africa.

Although current financial and human resources are not adequate to confront the mammoth challenge of curbing HIV/AIDS, the existing donor funds have stretched the absorptive capacity of many organizations to the limit. In Africa – especially in those countries in which as many as one third of the adults in urban areas are infected – many health workers have themselves been stricken. In short, available human resources are generally mobilized to implement programs; evaluation in this context of crisis understandably has taken a back seat.

However, with the epidemic showing no signs of letting up, international donor agencies realize that they must remain committed to HIV/AIDS initiatives for the long run. As the sums of money to combat HIV/AIDS increase, so do the calls for accountability in terms of the effective use of funds. The need for better monitoring and evaluation has also led to a growing number of data collection instruments and indicators.

The first global initiative to standardize evaluation indicators came during the 1980s in connection with the World Health Organization's Global Programme on AIDS (GPA). Experts from key organizations met in Geneva over a several year period in the late 1980s and early 1990s to develop a set of 10 "prevention indicators" (known as "PIs"). This initiative was highly laudable, in that standardization of indicators promised multiple benefits: countries with slim evaluation resources would not have to reinvent the wheel; rather, all could benefit from the guidance of this group of experts. Moreover, data from this standardized set of indicators would lend themselves to cross-national comparisons. However, the application of the PI in specific countries proved more challenging. Few countries were able to

produce the necessary data on the indicators, and/or the data were of questionable validity. Only a few countries used the PIs for monitoring of trends in programs, knowledge, or behaviors.

Over the last decade, much has changed in HIV/AIDS programming. Although prevention remains paramount, programs must also provide care and support of those infected and affected by the epidemic. Accordingly, evaluators have needed to adapt the monitoring and evaluation tools for these more complex programs.

M&E efforts in HIV/AIDS have evolved significantly in the past three years through a partnership that includes MEASURE *Evaluation*, UNAIDS, WHO, the Centers for Disease Control and Prevention (CDC), Family Health International (FHI), and other USAID cooperating agencies involved in HIV/AIDS programming and evaluation. The group first reviewed existing indicators and tools. Subsequently, MEASURE *Evaluation* and UNAIDS coordinated a series of case studies to review the HIV/AIDS M&E systems in 13 countries, including the feasibility of collecting data on the different PIs. These case studies provided a basis for (a) redefining the list of indicators most useful in monitoring and evaluating HIV/AIDS programs and (b) modifying or in some cases developing tools to measure the indicators.

In 2000, UNAIDS and its partners[1] published a manual entitled *National AIDS Programmes: A Guide to Monitoring and Evaluation*. We have reproduced the "core indicators" from the UNAIDS manual in this section of the *Compendium* as they appeared in the original, because this set of indicators was compiled through a very systematic and participatory process with input from experts throughout the world. The guide currently serves

---

[1] The partners included CDC, European Union, FHI/IMPACT, MEASURE DHS+, MEASURE *Evaluation*, SYNERGY Project, The POLICY Project, UNICEF, USAID, WHO, and the World Bank.

to strengthen M&E plans and implementation in many countries, as part of the expanded response against AIDS, especially in Africa.

In spite of the recent progress on M&E for HIV/AIDS, major methodological challenges remain.

Methodological Challenges of Evaluating STI/HIV/AIDS Programs

- **Measures requiring self-report are fraught with bias.**

This statement is true for survey work in general, but it is particularly applicable to HIV/AIDS for two reasons. First, respondents may not know their status (e.g., many women have asymptomatic sexually transmitted infections; few adults in the developing world know their HIV sero-status). Second, the stigma surrounding HIV/AIDS makes it a highly sensitive topic. Even if respondents knew their status, evaluators would have strong ethical issues in asking them to reveal it in the course of an interview. Surveys remain a fairly reliable means of determining levels of knowledge (e.g., on HIV transmission and prevention mechanisms). By contrast, questions regarding attitudes and sexual behavior may yield biased answers, influenced by communication programs that teach the "politically correct responses" regarding HIV/AIDS.

- **In contrast to family planning, HIV/AIDS lacks a set of reliable, easily measurable evaluation indicators.**

Comparing HIV/AIDS and family planning is useful to understand the dilemma facing those who evaluate HIV/AIDS programs. One might argue that the comparison is unfair, given that the FP research community has had an additional 20 years to wrestle with these methodological problems. Nonetheless, the problems are fundamental to the behaviors surrounding HIV/AIDS. (This infectious disease has a more complex web of proximate determinants and biological outcomes than does fertility.)

Much of FP evaluation has used three basic indicators:
- o Program level: Couple-years of protection (CYP)
- o Intermediate outcome: Contraceptive prevalence rate
- o Long-term outcome: Fertility rate

Regarding a **program level indicator**, the volume of condoms distributed for HIV/AIDS prevention is analogous to CYP. However, in the case of family planning, the evaluator can estimate the correspondence between the quantity distributed and the protection achieved (e.g., it takes four shots of Depo-Provera to cover a women for one year). By contrast, for condoms, such estimates of coverage – combined with the issues related to consistency of use – make it difficult to convert volume of condoms distributed to the number of persons protected against HIV/AIDS.

In terms of an **intermediate indicator** that measures a behaviour directly linked to the long-term outcome, condom use for HIV/AIDS is certainly analogous to contraceptive use. However, condom use carries a number of limitations. First, condom use is not the only means to safer sex (abstinence and reducing the number of partners lower the risk of HIV/AIDS). Second, the measurement of condom use in surveys is much less precise than the measurement of contraceptive prevalence (the latter which can be validated using fertility rates). In contrast to contraceptive use (which married couples generally practice over a period of time in a fairly consistent manner), condom use may be far more erratic and may vary by type of partner. Self-report on extramarital relations and condom use within such arrangements often lacks precision.

With respect to a **long-term outcome**, the objective of most HIV prevention programs is conceptually clear: to decrease the incidence rate of new HIV infections in a given population. However, it is virtually impossible in practical terms to obtain this information on a truly representative sample in a given country. To measure incidence would require identification and recruitment of a randomly selected sample of the population, followed across multiple (or at least two) periods of time. Moreover, any type of experimental design used to test the effects of a specific intervention is also fraught with the ethical difficulties of withholding a potentially valuable intervention from one group in the name of "good science." Other obstacles include the methodological difficulties in potential loss to follow-up, ethical issues related to lack of facilities to treat persons identified as sero-positive, and the high cost of such data collection. In fact, evaluators in this area have adopted a proxy approach: to test women attending antenatal clinics on the grounds that they represent a wide cross-section of the population. This indicator is the measure of choice for tracking the epidemic in many countries.

The international AIDS community now advocates second-generation surveillance, which involves coordinating the collection of several different types of data. Traditional surveillance systems tracked HIV infection or other biological markers of risk such as STIs. Because one of a limited number of behaviors – namely injection with contaminated needles or blood products or unprotected sex with an infected partner – must precede HIV/AIDS, we know that if these behaviors change, the spread of HIV will change. Second-generation surveillance systems monitor risk behaviors, using them to warn us of, or explain changes in, levels of infection. Thus, second-generation surveillance uses data from behavioral surveillance to interpret data gathered from sero-surveillance efforts.

- **HIV prevalence reflects the accumulation of HIV cases and changes slowly, even if incidence is on the decline.**

Obtaining data on HIV prevalence (e.g., from women attending prenatal clinics) is easier than obtaining data on HIV incidence, which requires tracking persons over time to determine the percentage of sero-negative individuals who become sero-positive in a given year. However, HIV prevalence has the limitation that the measure includes not only the cases (individuals) that become sero-positive in the past year, but rather in all previous years (and who are still alive). Thus, even if one

had an intervention that was successfully decreasing incidence of HIV in a given population, this trend would not be immediately evident in data on HIV prevalence. Evaluators have addressed this problem by focusing on 15-19 year olds attending antenatal care in a given country. In practical terms, none would have been sero-positive prior to initiation of sexual activity, which in much of Africa takes place around age 15. (Most female babies with perinatal AIDS would not have survived to adulthood.) Thus, time series data on the prevalence of HIV/AIDS among 15-24 year olds attending antenatal care serves as a useful proxy for HIV incidence. HIV prevalence trends among antenatal clients 15-19 may also be a good indicator of HIV incidence, but as simulations have shown, these trends are more sensitive to changes that may be unrelated with true incidence changes in the female population.

In sum, the indicators presented in this section are reproduced from the document entitled *National AIDS Programmes: A Guide to Monitoring and Evaluation*, which represents the joint efforts of UNAIDS and multiple partner organizations. We have modified the format used in their original document to be consistent with that used in this *Compendium,* and we have included only "core indicators" (not the alternative indicators). Otherwise, the content remains true to the original.

AIDS PROGRAM EFFORT INDEX (API)

Definition

The average score given to a national program by a defined group of knowledgeable individuals asked about progress in over 90 individual areas of programming, grouped into 10 major components

The API uses key informants from a designated mix of institutions to give opinions about central areas of commitment and programming, compiling an index out of scores given in various areas. The score, which is calculated as a percentage with 0 indicating no program effort and 100 indicating maximum effort, may be converted into a grade to minimize informant variation. Suggested grades range from very weak and weak through moderate and strong to very strong, depending on the range in which the numerical scores fall.

Data Requirement(s)

Responses to a detailed questionnaire from selected key informants (representatives of the Ministry of Health; NGOs; international consultants familiar with that country; and other informed individuals)

Data Source(s)

The AIDS Program Effort Index (API) questionnaire and protocol[2]

Purpose and Issues

The AIDS Program Effort Index is a composite index designed to measure political commitment and program effort in the areas of HIV prevention and care. It tries to capture many of the inputs and outputs of a national HIV/AIDS program. The score is made up of 10 main components of an effective national response:

- Political support;
- Policy formulation;
- Organizational structure;
- Program resources;
- Evaluation and research;
- Legal and regulatory aspects;
- Human rights;
- Prevention programs;
- Care programs; and
- Service availability.

The major concerns are that the API may be subjective and unreliable. The outcome depends entirely on the choice of informants, and informants will likely change from year to year. Because the indicator is still under development, the choice of informants has not yet been standardized.

Questions have also been raised about the utility of a single composite score, in which deterioration in some areas may mask improvements in other areas. For diagnostic as well as monitoring purposes, simply publishing the scores for the index separately for each component may be more useful. The separate category scores may stand alone as indicators, although for several areas of program effort, this document proposes alternatives based on measured parameters rather than on expert opinion and may therefore more usefully track trends over time.

One area in which the API process may yield a particularly useful indicator is in the area of policy formulation.

---

[2] Available online at: http://www.tfgi.com/api.asp

CONDOMS AVAILABLE FOR DISTRIBUTION NATIONWIDE

Definition

The total number of condoms available for distribution nationwide during the preceding 12 months, divided by the total population aged 15-49

The indicator estimates the number of condoms (male and female) available for in-country use during the last 12 months. Key informants are identified and interviewed to uncover all possible sources of condom manufacture, import, distribution, and storage. Next, data are collected from all manufacturers and major commercial distributors as well as major donors, condom storage facilities, and government, parastatal, and NGO bodies involved in acquiring and distributing condoms.

This indicator sums the condoms in stock nationally at the start of the 12-month period, plus condoms imported during the 12-month period, plus condoms manufactured in country during the same period, minus any exports of condoms over that period. The sum of all condoms available for use in the country during the past 12 months is then divided by the total population aged 15-49.

This indicator is calculated as:

$$\frac{\text{\# of condoms in stock} + \text{\# of condoms imported} + \text{\# of condoms manufactured} - \text{\# of condoms exported during a 12-month period}}{\text{Total population aged 15-49}}$$

Data Requirement(s)

Condom inventories and purchase records for (1) social marketing programs, (2) national condom distribution programs, and (3) commercial wholesalers/distributors; stock records and production records for manufacturers

Data Source(s)

WHO/GPA protocol for estimating condom availability for distribution at the central and peripheral level

Purpose and Issues

The best distribution system in the world is not much help if it has nothing to distribute. The first challenge for national programs promoting condom use is to ensure that there are enough condoms in the country to satisfy demand. This indicator measures the number of condoms available for use by those in the most sexually active age group. Where programs actively promote the availability of male condoms, they should also include female condoms, although the indicator should be disaggregated by condom type.

Combining this indicator with indicators of sexual behavior gives a powerful picture of the adequacy of condom provision. For example, if a third of all men aged 15-49 say they have had non-regular sex in the past year, and 20 percent of married couples say they have used condoms to avoid pregnancy, and yet only three condoms are available per sexually active adult per year, one can deduce that the national supply of condoms is insufficient to meet the potential demand.

The number of condoms available at the central level helps assess the adequacy of overall condom availability. We must note, however, that "availability" does not equal "accessibility," which includes dimensions of price, location, and access by sub-populations at risk for unprotected sex and HIV. Frequently, not all available condoms are distributed or reach the individuals who most need them to protect against the spread of HIV. This indicator by itself cannot give a picture of how many "in-stock" condoms actually get distributed or used.

Ironically, efforts at the national level to encourage condom use sometimes complicate the measurement of this indicator. Many countries have deregulated condom imports in the face of AIDS in order to maximize the number of condoms available. Deregulation means that a wide variety of companies, NGOs, donors and gov-

ernment departments (the health ministry, the defense ministry, among others) may import condoms without necessarily reporting numbers imported to a central body. Programs distinguish between condoms distributed through family planning programs and those distributed to reduce sexually transmitted infections. One must take both sources into account. If possible, evaluators need to present data by program, because family planning programs primarily intend condoms for relatively low-risk acts within stable monogamous unions, whereas AIDS program condoms aim at higher risk sexual contacts.

Where condom promotion activities center around marketing condoms at subsidized prices to people likely engaging in risky sex (social marketing), sales of particular brands of condoms can also provide a useful indicator of program success. Organizations responsible for the social marketing of condoms typically keep very good records of condoms distributed down to the retail level. Although these data tell only part of the story of condom availability, they provide a very low-cost source of information for the National AIDS Program and can be very useful for advocacy purposes. A rise in the number of condoms manufactured or imported into a country, or of condoms sold, can be useful in supporting other indicators measuring rises in self-reported condom use or falls in self-reported STIs and, eventually, in HIV prevalence.

PERCENT OF POPULATION WITH ACCEPTING ATTITUDES TOWARDS THOSE LIVING WITH HIV

Definition

The percent of people who express accepting attitudes towards people with HIV, of all people surveyed aged 15-49

Respondents in a general population survey answer a series of questions about people with HIV, as follows:

- If a member of your family became sick with the AIDS virus, would you be willing to care for him or her in your household?
- If you knew that a shopkeeper or food seller had the AIDS virus, would you buy fresh vegetables from them?
- If a female teacher has the AIDS virus but is not sick, should she be allowed to continue teaching in school?
- If a member of your family became infected with the AIDS virus, would you want it to remain a secret?

Only a respondent who reports an accepting or supportive attitude on all four of these questions enters the numerator. The denominator is all people surveyed.

This indicator is calculated as:

$$\frac{\text{\# of respondents who report accepting attitudes on all four questions}}{\text{Total \# of respondents aged 15-49}} \times 100$$

Data Requirement(s)

Self-reported data from survey respondents

Data Source(s)

UNAIDS general population survey; DHS AIDS Module; FHI BSS (adult); FHI BSS (youth)

Purpose and Issues

This indicator is based on answers to a series of hypothetical questions about men and women with HIV. It reflects what people are prepared to say they feel or would do when confronted with various situations involving people living with HIV.

Methodologically, this is a relatively easy way to construct an indicator of attitudes to people with HIV. A low score on the indicator is a fairly sound indication of high levels of stigma, and for that reason alone it is worth measuring.

However, difficulties arise in interpreting indicators based on hypothetical questions, and a high score on the indicator is harder to understand than a low one is. It could mean little real stigma is attached to HIV. Or it could mean that people know they should not discriminate, and therefore report accepting attitudes. This knowledge may not change their behavior, which may continue to be discriminatory towards people with HIV. Changes in the indicator may therefore reflect a reduction in stigma or simply a growing awareness that it is not nice to own up to one's prejudices. That awareness in itself may, however, constitute the first step in program success. High scores may also reflect the respondent's limited personal experience with someone who is HIV-infected.

The proposed indicator is similar to an earlier measure developed by WHO, but questions have been changed following field testing to better reflect situations in which people with HIV actually suffer from stigma. Field tests revealed that the exact wording of the indicator greatly affected responses. When the gender of the teacher was not specified, for example, one country registered very high levels of "discriminatory" attitudes on that question, for example. Further investigation showed that the negative attitudes were related to recent news reports of male teachers infecting female pupils with HIV.

The earlier WHO indicator has been little used, calling into question the utility of this measure. Possibly, it was little used because so little programming effort to date has gone in to reducing stigma surrounding HIV in most countries. As the power of stigma to obstruct preven-

tion and care efforts becomes ever clearer, however, more national AIDS programs will likely turn their attention to this area. We expect, therefore, that use of this indicator will increase.

Some have suggested that this indicator serves to measure differences in discrimination or stigma by gender. Although some research suggests that women are more likely than men to be treated and viewed harshly if they have HIV or AIDS, other recent surveys have shown little difference in response to gender-specific questions about stigma and discrimination.

PERCENT OF POPULATION WHO KNOW HIV PREVENTION METHODS

Definition

The percent of all respondents who, in response to prompted questions, say that a person can reduce the risk of contracting HIV by using condoms or having sex only with one faithful, uninfected partner.

The denominator includes all respondents surveyed, regardless of whether they have ever heard of AIDS or not. The indicator components should also be reported separately to show changes in specific knowledge areas.

The evaluator (i.e., the person designing the questionnaire) must carefully select the precise wording of the prompted questions in each linguistic and cultural context. We should note that the correct prevention methods prompted for should be interspersed in the questionnaire with misconceptions used to calculate the following indicator: **Percent of Population with No Incorrect Beliefs About AIDS**.

This indicator is calculated as:

$$\frac{\text{\# of respondents who say that a person can reduce the risk of contracting HIV by using condoms or having sex only with one faithful, uninfected partner}}{\text{Total \# of respondents}} \times 100$$

Data Requirement(s)

Self-reported data from survey respondents

Data Source(s)

UNAIDS general population survey; DHS AIDS module; FHI BSS (adult); FHI BSS (youth)

Purpose and Issues

Most AIDS programs targeting the general population promote mutual monogamy and condom use as the primary ways of avoiding HIV infection among the sexually active men and women who make up the majority of all adults in virtually every population. This indica-

tor measures the extent to which those messages have reached the general population or the specific sub-population surveyed.

Data for this indicator are easy to collect in a population survey. In most countries, the score on this indicator will be high, but disaggregation of the indicator by individual questions, residence, gender, or age group may provide useful pointers to gaps in information flows.

Limitations of the use of prompted data were discussed in the introduction to this section. Although the evaluator should construct the primary indicator using prompted data, a comparison between prompted and non-prompted data where possible may yield interesting information. For instance, both the revised UNAIDS general population survey and the DHS AIDS module ask, "What ways can people protect themselves from getting HIV?" before asking specific "prompted" questions. To further help program managers, this indicator should always be used in conjunction with: **Percent of Population with No Incorrect Beliefs about AIDS**

Previous knowledge indicators included abstinence as a "correct" method of prevention used in this indicator. Abstinence is an extremely important prevention option for young people. However, research in many settings shows that adults who are already sexually active rarely use abstinence as a primary HIV prevention method. In addition, people who know that HIV is sexually transmitted are highly likely to know that not having sex can reduce the risk of transmission. Negative responses on this item are more likely to result from people believing that abstinence is not feasible than from their believing that abstinence does not provide effective protection. In surveys among young people, however, questions about abstinence continue to be important. Programs focusing on delaying age at first sex among young people may choose to add a knowledge indicator that includes correct responses to a question about abstinence as a prevention method.

Gender Implications of this Indicator

Women may be poorly informed about prevention methods because of gender-related socio-cultural barriers that deny or discourage their access to information about HIV and their willingness to discuss such information. For example, many cultures view a woman's ignorance about sexual matters as a sign of purity, this being an ideal feminine quality. Women may thus be reluctant to seek out information about HIV prevention for fear of being labeled as promiscuous (UNAIDS, 1999a). Additionally, women may be reluctant to report that they have such knowledge, for the same reason.

HIV prevention messages may also not be reaching women because some providers may be reluctant or fail to discuss condom use as a method of preventing HIV with female clients. They may perceive condoms as a male contraceptive or because they may associate condoms with illicit sexual activity.

For example, in areas where female literacy rates are low, information campaigns on HIV prevention that rely on printed materials may be ineffective in reaching women who are illiterate or have limited literacy skills.

Percent of Population with No Incorrect Beliefs about AIDS

Definition
The percent of all respondents who correctly reject the two most common local misconceptions about AIDS transmission or prevention, and who know that a healthy-looking person can transmit AIDS

In a series of prompted questions, the interviewer reads a series of correct and incorrect statements about AIDS transmission and prevention. Responses to the correct statements about prevention are used to calculate the previous indicator: **Percent of Population Who Know HIV Prevention Methods**. Responses to a question about infection status in healthy-looking people and to two incorrect statements about transmission or prevention are used to calculate this indicator.

The incorrect statements will vary to reflect the misconceptions most common in the local context. Very often these misconceptions will include the belief that AIDS can be spread through an insect bite or through witchcraft. Sometimes, they will include beliefs about prevention or cure, such as AIDS being preventable by eating certain types of foods or herbs, or being curable by having sex with a certain type of person such as a virgin (or simply being curable at all). One question will always center on knowledge of the "healthy carrier" concept, that is, knowledge that a person may contract HIV by having unprotected sex with an apparently healthy person. The exact wording may vary locally. For example, in some areas "fat" may be synonymous with "healthy" in this context and may better reflect people's misunderstanding of who constitutes a "safe" partner.

Before conducting a survey, the evaluator should identify the local misconceptions. They may vary over time within the same country.

To enter the numerator for this indicator, a respondent must correctly reject both misconceptions, and must know that a healthy-looking person can transmit AIDS. The denominator is all respondents, including those who have not heard of AIDS. For program purposes, the in-

dicator should be disaggregated by misconception, and the percentage believing that a healthy-looking person cannot transmit HIV should also be reported separately.

This indicator is calculated as:

$$
\frac{\substack{\text{\# of respondents who correctly reject the}\\ \text{two most common local misconceptions}\\ \text{about AIDS and who know that a healthy-}\\ \text{looking person can transmit AIDS}}}{\text{Total \# of respondents}} \times 100
$$

Data Requirement(s)
Self-reported data from survey respondents

Data Source(s)
UNAIDS general population survey; DHS AIDS module; FHI BSS (adult); FHI BSS (youth)

Purpose and Issues
Many of the people who know that condoms protect against AIDS also believe that AIDS can be contracted from a mosquito bite or other uncontrollable event. Why bother to reduce the pleasure of sex, they reason, if they might in any case be infected by something as random as a mosquito bite? At high levels of HIV-related awareness, a reduction in misconceptions that act as a disincentive to behavior change may actually better reflect the success of a BCC campaign than will an incremental shift in already high levels of "correct" knowledge. This indicator measures progress made in reducing misconceptions.

Again, this indicator is easy to measure. It gives a good picture of the level of false beliefs that may impede people's determination to act on correct knowledge. When the data are disaggregated, they provide invaluable information for program managers planning future BCC campaigns, telling them which misconceptions must be attacked, and in which sub-populations.

A word of caution is in order, however. There is always a danger that including misconceptions in a question-

naire actually increases their credibility. Preparatory research should be sure to establish commonly held misconceptions (rather than run the risk of promoting new ones), and the questionnaire should make very clear that some of the statements in the sequence are true while others are false.

One limitation is the indicator's potential inability to distinguish between misconceptions which are likely to influence behavior and those which are merely incidental. Measurement of this indicator also requires preparatory work to determine which misconceptions are currently most likely to be common.

PERCENT OF POPULATION WHO KNOW METHODS OF PREVENTING MOTHER-TO-CHILD TRANSMISSION OF HIV

Definition

The percent of women and men who correctly respond to prompted questions about preventing mother-to-child transmission of HIV through anti-retroviral therapy and avoidance of breastfeeding

Respondents in a population survey answer a series of questions about the transmission and prevention of HIV. (See previous indicators: **Percent of Population Who Know HIV Prevention Methods** and **Percent of Population with No Incorrect Beliefs about AIDS**.) Among these are questions about whether HIV can be transmitted from mother to child and about means of preventing mother to child transmission.

The indicator is the number of respondents who say that HIV transmission from women who have tested HIV positive can be prevented by the mother taking drugs during pregnancy, and by the mother avoiding breastfeeding, divided by the total number of respondents to the survey.

This indicator is calculated as:

$$
\frac{\text{\# of respondents who say that a mother can avoid MTCT of HIV by taking drugs during pregnancy and by avoiding breastfeeding}}{\text{Total \# of respondents}} \times 100
$$

Data Requirement(s)

Self-reported data from survey respondents

Data Source(s)

UNAIDS general population survey; DHS AIDS module; FHI BSS (adult)

Purpose and Issues

This indicator examines whether women and men know of methods to prevent the transmission of HIV from mother to child. In this field, as in the field of prevention of sexual transmission, knowledge is a prerequi-

site for decision-making and intervention, although by no means sufficient to ensure it.

This indicator measures people's knowledge of methods to prevent transmission from mother to child through anti-retroviral therapy and avoidance of breastfeeding. Men's knowledge in this area is also important, not least because in many societies men dominate decisions about family formation and childbearing, so the indicator is constructed for both sexes. Because most BCC campaigns in this area aim to reach women, program managers will want to monitor program effectiveness by disaggregating the indicator by gender.

This indicator presupposes that efforts are being made to educate women about maternal to child transmission of HIV and that information about prevention forms part of that education.

The indicator does not distinguish in its denominator between those who know about maternal to child transmission and those who do not, because people who do not know that such transmission can be prevented are definitely among those who have not been reached with information about prevention methods. The questioning sequence does, however, allow countries to construct an indicator of knowledge about HIV transmission from mother to child should they wish.

The knowledge that transmission from mother to child can be prevented is likely to shape women's care-seeking and breastfeeding behavior. A pregnant woman who simply knows that she can pass HIV on to her child is less likely to seek to know her HIV status than a pregnant woman who knows that she can avoid transmitting HIV to her child.

In many countries in Latin America and elsewhere, the demand for prevention has driven a radical improvement in service provision for pregnant women with HIV. Such a demand will only arise if people know that therapy exists and can effectively reduce transmission of HIV to infants.

PERCENT OF POPULATION REQUESTING AN HIV TEST, RECEIVING A TEST, AND RECEIVING TEST RESULTS

Definition

The percent of people aged 15-49 surveyed who have *ever* voluntarily requested an HIV test, have received the test, and have received their results

The evaluator can also collect data on those requesting an HIV test, receiving the test, and receiving their results *in the last 12 months.*

A general population or sub-population survey can include questions on whether the respondent has ever requested an HIV test, whether he/she has been tested, and if so, whether he/she has received the results. Those having ever requested a test and having received the results form the numerator, while the denominator is all respondents in the survey.

The interviewer prefaces the question by saying, "I do not want to know the results of the test…." As for most indicators, results should be presented by component and separately for men and women. In addition to having information on the broad reach of VCT services over time, knowing the percentage of the population surveyed who have been tested and have received the results *in the last 12 months*, a more time-sensitive measure, will be useful.

This indicator is calculated as:

$$\frac{\text{\# of respondents aged 15-49 who requested, underwent, and received HIV test results}}{\text{Total \# of respondents aged 15-49}} \times 100$$

Data Requirement(s)

Self-reported data from survey respondents

Data Source(s)

UNAIDS general population survey; DHS AIDS module; FHI BSS (adult); FHI BSS (youth)

Purpose and Issues

The coverage of quality VCT services will go a long way towards determining whether those services achieve their threefold aims of providing an entry point for care and support, promoting safe behavior, and breaking the vicious circle of silence and stigma.

This indicator aims to estimate the reach of HIV testing services in the general population and of the percentage of people who now know their HIV status. It can also be used in specific sub-populations with high-risk behavior where counseling and testing services are being promoted. When calculated for sub-populations with high-risk behavior, the numerator should include only those who requested a test and received their results in the last 12 months.

A breakdown of the indicator into its components parts (looking, for example, at people who requested and received a test but never received their results) can point to gaps in program service provision and quality of care. Data on those who do not return for results or do not know their results may offer insight, for example, into levels of stigma and/or reluctance to learn HIV status given the lack of available options for care.

The survey question for this indicator specifies that the respondent must have requested the test. In many situations, people may assume that their blood has been tested for HIV at some time, for example, when giving a blood donation, when applying for insurance, or for surveillance purposes when attending antenatal services. To calculate this indicator, the evaluator excludes these involuntary tests, whether real or perceived. Also excluded are tests made for diagnostic purposes without the consent of the client, even if the client was later told of the results. Such tests do not reflect either the coverage of or the demand for testing services, nor do they take into account that the measure emphasizes the "voluntary" element desired for HIV tests. For that reason, survey questions must specify that the person requested a test.

In many countries, many people will have been offered and accepted an HIV test in a health care setting. To measure the proportion of people who may be aware of their sero-status (regardless of who initiated the request for a test), one should also collect data on people having been offered a test, having accepted it, and having received their results.

This indicator gives some idea of the increasing coverage of services that meet people's demand for testing. It is not, however, limited to voluntary testing and counseling services staffed by trained counselors. It may, therefore, include tests requested from private doctors who do not necessarily provide any counseling.

In areas where HIV is highly stigmatized, respondents may be unwilling even to admit to having taken an HIV test, because such an admission may imply that they fear they may be infected. This unwillingness increases when the question arises in a questionnaire about risk behavior. On the other hand, in countries that heavily promote testing as a "responsible" act, some people may say they have been tested when in fact they have not. Despite these potential biases, the indicator is useful for providing a rough estimate of the proportion of people likely to know their HIV status at all.

If one adapts the indicator to reflect the percentage of respondents requesting, receiving an HIV test and receiving results *in the last 12 months*, the measure will reflect recent changes in testing services, knowledge about testing among the population surveyed and desire for testing. While those people exposed to HIV more than once in a lifetime should be targeted for repeat testing, existing evidence suggests that most only get tested once. Thus, in high prevalence populations with good coverage of testing, those who tested positive in the past may not return for further testing in future years. This could affect trends in the time-bound indicator. The tendency to test only once tends to bias this indicator in underestimating prevalence.

The "ever tested" measure is less sensitive to recent trends in test-seeking behavior than a time-bound measure such as "tested in the last 12 months," but the "ever-tested" measure will provide an idea of the overall reach of testing services.

In low-level and concentrated epidemics, the indicator will likely yield extremely low percentages if measured in the general population. However, the indicator is appropriate to measure behavior in sub-populations at higher risk of infection.

Gender Implications of this Indicator

Women's use of VCT services may be limited because of a variety of barriers that prevent them from accessing services, such as: lack of mobility, the cost of tests, and reluctance or prohibition of receiving care from a male health care provider. A lack of female health care providers or female-oriented health services may prevent women from requesting VCT. Women, especially those who are married or who believe they are in a monogamous relationship, may have a low perceived risk of being infected with HIV and thus may not request VCT. Research suggests that men are more likely to seek out VCT services without consulting others, while women feel obligated to discuss testing with partners before requesting services, thus creating a potential barrier to VCT access (Gupta, 2000).

A positive test result may have serious repercussions for women, including partner violence, isolation, ostracism, or abandonment. For this reason, women, although they suspect they may be infected, may not request VCT or return for results. HIV positive women may also fail to fully comply with recommended treatments or bottle-feeding to prevent MTCT for fear of being identified as HIV positive (UNAIDS, 1999b).

VOLUNTARY COUNSELING AND TESTING CENTERS WITH MINIMUM CONDITIONS TO PROVIDE QUALITY SERVICES

Definition

The number and percent of facilities that provide VCT services meeting minimum conditions necessary to provide quality counseling and HIV testing services

Evaluators take a random sample of facilities and assess the quality of counseling and testing services (including NGOs, private clinics and doctors' surgeries). They evaluate structural elements necessary to provide quality counseling and testing services, including trained staff, adequate privacy for counseling, systems for maintaining confidentiality, a directory of services for referral, and adequate conditions for ensuring quality control of specimen tests.

Evaluators may choose to weight the score obtained by each site in the random sample by the annual client load of that site. The indicator is the number of clients served in the last year by sites with adequate conditions to provide quality VCT services, divided by the total number of clients served in the last year by all sites sampled.

This indicator is calculated as:

$$\frac{\text{\# of clients served in the last year by sites with adequate conditions to provide quality VCT services}}{\text{Total \# of clients served in the last year by all sites sampled}} \times 100$$

Data Requirement(s)

Assessment by external evaluator of adequacy of key elements for conducting VCT

Data Source(s)

UNAIDS protocol for the evaluation of voluntary counseling and HIV testing services

Purpose and Issues

In many countries, voluntary counseling and testing has landed in the hands of under-funded and ill-equipped non-government and community organizations or has become a corollary of private sector health service providers. Many of these entities lack even the most basic structural facilities necessary to provide quality counseling, such as a room where counseling can occur privately, or a regular electricity supply to ensure the adequate storage of specimens until testing.

This indicator measures a condition that is necessary but not sufficient to guarantee quality counseling services. The percentage of clients served in a facility that meets conditions for quality counseling is also likely to reflect other factors, such as access, available testing services, or the history of positive experiences at the center by other community members. Inevitably, a number of contextual variables influence the results of an indicator assessing quality. The goal of the indicator is to provide a framework for assessing some accepted goals and guidelines.

A potential difficulty in constructing this indicator is that sites with inadequate record keeping may be unaware of their overall client load; it will therefore be impossible to weight the indicator by client load. One could construct the indicator as a simple percentage (i.e., the percentage of facilities surveyed which meet minimum conditions for adequate service). However, because poor conditions at a small facility with a low caseload is relatively less important than poor facilities at a large and busy center, one should apply weighting where possible. (In truth, a strong correlation may exist between conditions and caseload: caseloads may be low *because* conditions are poor.)

As with other aggregate indicators, the evaluator may need to obtain information on different elements separately for program planning purposes. Disaggregating this indicator by type of service provider (NGO, hospital, private clinic) may also be useful.

PERCENT OF PREGNANT WOMEN COUNSELED AND TESTED FOR HIV

Definition

The percent of women who received counseling during antenatal care for their most recent pregnancy, who accepted an offer of testing, and who received their test results, based on all women pregnant at any time in the two years preceding the survey

In a general population survey, the interviewer asks the respondent when her most recent child was born and whether she received any antenatal care before that last birth. If she received care, she is asked whether clinic staff talked to her about HIV infection and offered her a confidential HIV test. If yes, she is further asked if she agreed to a test and if she received the results. Before asking such questions, the interviewer offers assurance that he/she is not interested in knowing the outcome of any test.

The indicator is the number of women counseled and offered voluntary HIV testing at ANC before their most recent birth in the last two years and received their test results, divided by the total number of women surveyed. To measure recent trends the evaluator excludes women whose most recent birth was more than two years ago from the analysis.

This indicator is calculated as:

$$\frac{\text{\# of women who were counseled, who accepted the offered voluntary HIV testing at ANC, and received results before their most recent birth in the last two years}}{\text{Total \# of women who were pregnant in the last two years}} \times 100$$

Data Requirement(s)

Self-reported data from survey respondents

Data Source(s)

UNAIDS general population survey

Purpose and Issues

The principal active interventions to reduce mother to child infection depend on knowledge of HIV status. Knowledge of HIV status during pregnancy may also affect future reproductive choices. Ideally, women would learn their HIV status using VCT services before they choose to become pregnant. But the gap between this ideal and reality is often very wide. In practice, the first opportunity many women have to be counseled about HIV and to be offered tests may be at antenatal clinics that offer these services as a precursor to offering interventions to reduce transmission of HIV from mother to child.

To learn their HIV status in an antenatal care situation, women have to complete a number of steps. First, they must attend antenatal services. Then, they must be counseled and offered an HIV test. Next, they must accept a test. Finally, they must return to receive the test results. It is only after the post-test counseling that follows all of these steps that they make necessary decisions about therapy and infant feeding.

This indicator measures the percentage of women with a recent pregnancy who completed all of those steps. A general population survey can yield data on the coverage of ANC-based counseling and testing country-wide, rather than just in specific pilot facilities.

This broad measure of service provision reflects coverage on a national scale. However, few countries may have the resources to introduce counseling and voluntary testing for pregnant women country-wide. Those countries providing prevention services for pregnant, HIV-positive women typically start with pilot projects in a few antenatal clinics. Even if all women in pilot clinics are counseled and offered testing, the indicator will typically remain low for some time. Evaluators should use this indicator in conjunction with the following one: **Antenatal Clinics Offering and Referring for VCT**.

As a summary indicator, it does not attempt to diagnose at which point women are dropping out of the spectrum of care. For program purposes, managers will find it important to know whether a poor result on the summary indicator is because of low initial attendance at antenatal services, because women attending services are not being offered tests, because they are refusing the offer of a test, or because they are tested but do not return for test results. Each of these points of failure has a different implication for programming, and all can be calculated from the data collected for this indicator. The summary indicator does not attempt to measure quality of counseling or other elements of service coverage.

Gender Implications of this Indicator

The availability of anti-retrovirals which can prevent the transmission of HIV from mother to child increase the value that voluntary counseling and testing could have for pregnant women. Yet many women refuse testing and treatment. Health workers must recognize that VCT entails significant risks for women which may include partner violence and ostracism. Applying strict standards of confidentiality and privacy to VCT, as well as to the treatment phase (if one is required), is necessary to ensure that pregnant women will have enough trust in their own safety to risk being tested. The lack of treatment options for the mother herself remains a serious obstacle to prevention of MTCT.

ANTENATAL CLINICS OFFERING AND REFERRING FOR VCT

Definition

The percent of public antenatal clinics offering counseling and voluntary testing for HIV by trained staff or referring to VCT services

Evaluators randomly select public antenatal clinics in a health facility survey. Evaluators conduct staff interviews and record reviews to ascertain whether any of the clinic staff are trained in counseling, and whether the clinic routinely counsels clients about HIV in pregnancy and offers HIV tests with post-test counseling or refers clients to qualified outside services. The annual client volume of the clinic is also recorded.

This indicator is calculated as:

$$\frac{\text{\# of antenatal clinics offering voluntary testing for HIV and post-test counseling by trained staff}}{\text{Total \# of antenatal clinics}} \times 100$$

Evaluators may then weight the result by client volume.

Data Requirement(s)

Total number of antenatal clinics and number of antenatal clinics offering and referring for VCT

Data Source(s)

UNAIDS guide to the monitoring and evaluation of prevention programs for mother to child HIV transmission; UNAIDS tool for evaluating HIV voluntary counseling and testing

Purpose and Issues

Private sector clinics will often take the lead in providing services for those HIV-infected pregnant women who can afford to pay for interventions. Because such interventions are relatively expensive, the goal of national programs is to extend their reach to less affluent members of society, through service provision in public facilities. Thus, evaluators should calculate this indicator based on service provision in public sector clinics. However, countries that are making an effort to increase training in counseling for staff at antenatal clinics in the private sector or among traditional birth attendants may want to include such groups in this indicator as well.

Ideally, this measure would include all public antenatal services in a country. Since the number of such services is often too large to be practical, sampling is adopted.

This indicator is most useful in countries actively expanding coverage of maternal to child prevention services. A steady rise in the indicator is likely to reflect a steady expansion of service provision. However, if sampling is necessary, the indicator may be slow to reflect progress.

PERCENT OF POPULATION WHO HAD HIGHER RISK SEX IN THE LAST YEAR

Definition
The proportion of respondents who have had sex with a non-marital, non-cohabiting partner in the last 12 months of all respondents reporting sexual activity in the last 12 months

Respondents are asked about their marital status and the last three sexual partners within the last 12 months. For each partner, details are taken of cohabiting status as well as duration of the relationship, condom use, and other factors. The numerator is those respondents who say that in the last 12 months, they have had sex with someone who is not their spouse or the person they live with. The denominator for this indicator is all respondents who report having any sex in the last 12 months.

The numerator should exclude polygynous men who live with several wives unless they also have sex with women who are not part of their household.

This indicator is calculated as:

$$\frac{\text{\# of respondents who report having sex with a non-marital, non-cohabiting partner in the last 12 months}}{\text{Total \# of respondents who report having sex in the last 12 months}} \times 100$$

Data Requirement(s)
Self-reported data from survey respondents

Data Source(s)
UNAIDS general population survey; DHS AIDS Module; FHI BSS (adult)

Purpose and Issues
The spread of HIV depends upon unprotected sex with people who also have other partners. Most monogamous relationships are cohabiting, although the reverse is not necessarily true. Partners who do not live together – who have sex only occasionally – are those who are most likely to have other partners over the course of a year. These partnerships therefore carry a higher risk of HIV transmission than partnerships that remain outside a wider sexual network. AIDS prevention programs try to discourage high numbers of partnerships and to encourage mutual monogamy. This indicator estimates the proportion of the population that engages in relatively high-risk partnerships and that is therefore more likely to be exposed to sexual networks within which HIV can circulate.

This indicator gives a picture of levels of non-monogamous sex. If people stop having sex with all of their extramarital partners, the change will be captured by changes in this indicator. However, if people simply decrease from seven extra-marital partners to one, for example, the indicator will not reflect a change, even though this reduction may have a significant impact on the epidemic spread of HIV and may count as a program success.

This indicator proposes a different definition for higher risk sex than that commonly used in the past. Obviously, a change in definition will upset trend data for countries that have collected data using a different definition from the one currently proposed. However, this difficulty is surmountable. The proposed data collection instrument allows for both the old and new versions of the indicator to be calculated simultaneously. In practice, in existing data, which allow for the comparison between the two indicators, the difference has been small. The change is proposed largely because countries report dissatisfaction with previous indicators, arising mostly from respondents' difficulties in understanding the definitions of regular and non-regular partnerships.

PERCENT USING CONDOMS AT LAST HIGHER RISK SEX

Definition

The percent of respondents who say they used a condom the last time they had sex with a non-marital, non-cohabiting partner, of those who have had sex with such a partner in the last 12 months

For each partner listed in the last 12 months, respondents are asked whether they used a condom the last time the couple had sex. Other questions allow for the classification of partnerships as cohabiting or non-cohabiting. All those who report at least one non-marital, non-cohabiting partner in the last 12 months (i.e., the numerator of the previous indicator: **Percent of Population Who Had Higher Risk Sex in the Last Year**) form the denominator. The numerator is the number of those in the denominator who used a condom the last time they had sex with their *most recent* non-cohabiting partner.

This indicator is calculated as:

$$\frac{\text{\# of respondents who report using a condom the last time they had sex with a non-marital, non-cohabiting partner}}{\text{Total \# of respondents who report having sex in the last 12 months with a non-marital, non-cohabiting partner}} \times 100$$

Data Requirement(s)

Self-reported data from survey respondents

Data Source(s)

UNAIDS general population survey; DHS AIDS Module; FHI BSS (adult)

Purpose and Issues

If everyone used condoms every time they had sex with a non-marital or non-cohabiting partner, a heterosexually transmitted HIV epidemic would be almost impossible to sustain. Although AIDS programs try to reduce casual partnerships, they must also, if they are to successfully curb the epidemic, promote condom use in the casual partnerships that remain. This indicator tracks changes in condom use in these partnerships.

A rise in this indicator is an extremely powerful indication that condom promotion campaigns are having the desired effect among their principal target market.

Because condom promotion campaigns aim for consistent use of condoms with non-regular partners rather than simply occasional use, some surveys have tried to ask directly about consistent use, often using an always/sometimes/never question. Although this question may be useful in sub-population surveys (see below), it is subject to recall bias and other biases and is not sufficiently robust for use in a general population survey. Asking about the most recent act of non-cohabiting sex minimizes recall bias and gives a good cross-sectional picture of levels of condom use. However, since higher levels of condom use with non-regular partners will reflect increases in overall consistency of use, this indicator supports the objective of consistent condom use.

Gender Implications of this Indicator

Although women may know the protective effect of condoms, sexual negotiation between partners depends on the balance of power between partners, which in most places, weighs more heavily in the man's favor. This has several ramifications. Many women may lack the negotiation skills to ask their partner to use a condom or they may be reluctant to approach the subject because of the association between condoms, illicit sex, and STIs. Some women may be apprehensive about demanding or negotiating condom use (or withholding sex if partners refuse to use condoms) for fear of partner violence, fear of being perceived as unfaithful or promiscuous, or fear of abandonment (which some women may perceive as having more serious consequences than engaging in unprotected sex) [UNAIDS 1998].

Gender Issues (continued)

Many women may not be having "higher risk sex" (defined as having sex with a non-marital, non-cohabitating partner), but they may be exposed to HIV infection within a monogamous relationship or a marriage, especially where condom use is rare between marital or regular partners. Cultural norms are lenient with regards to men's multiple sex partners – thus diminishing the protective effect that a monogamous relationship has for women who are unable to control or are unaware of their partner's extra-marital relationships (UNAIDS, 1999a).

Indicator

PERCENT OF MEN HAVING COMMERCIAL SEX IN LAST YEAR

Definition

The percent of men reporting sex with a sex worker in the last 12 months

This indicator is intended *only* for countries with well-defined populations of sex workers. (See below.) In general population surveys or in specialized surveys among groups of men who fit the profile of clients of sex workers (the military, truck drivers, among others), men are asked directly if they had sex with a sex worker in the previous 12 months.

Although a given country may have several different types of definable sex workers, each with different perceived levels of risk, evaluators can combine all these groups into an indicator of commercial sex use for monitoring and evaluation purposes.

This indicator is calculated as:

$$\frac{\text{\# of men who report having sex with a sex worker in the past 12 months}}{\text{Total \# of male respondents}} \times 100$$

In some countries, evaluators have collected this indicator in the past using only sexually active men (rather than all male respondents) as the denominator. To maintain valid trend data, we recommend calculating this indicator both ways (using only sexually active men and using all men in the denominator) over a period of several years.

Data Requirement(s)

Self-reported data from survey respondents

Data Source(s)

UNAIDS general population survey; DHS AIDS Module; FHI BSS (adult)

Purpose and Issues

In concentrated epidemics, sexual mixing between groups with a high likelihood of infection and the general population is of central interest. In heterosexual concentrated epidemics, the initial focal point of infection is among sex workers and their clients. Those clients then spread infection to their wives and girlfriends in the general population, as well as to other sex workers. In such situations, AIDS programs often focus on trying to reduce the proportion of men having sex with sex workers, as well as increasing condom use in these encounters. This indicator measures progress towards the first of these goals.

This indicator is useful in concentrated heterosexual epidemics in countries where commercial sex (and especially brothel-based sex) is common, and where a "prostitute" has a clearly defined role. Thus the indicator is most likely to be used in parts of the world where commercial sex has played a dominant role in the epidemiology of HIV (e.g., many countries in Asia).

Attempts to collect and analyze data using a wider definition of commercial sex (questions such as "Have you given or received money or gifts in exchange for sex?") have not yielded useful information. In epidemic terms, sex workers are of interest because they have a high turnover of partners and therefore have a high probability of being exposed to infection and passing on infection. In many cultures, the high number of partners is true of only a fraction of the people who have "received money or gifts in exchange for sex." If no locally specific term for prostitution exists, the chances are that this indicator is not relevant to the program and should not be used.

The indicator is also of limited use in very high prevalence epidemics, because differences in risk associated with sex with a sex worker compared with any other casual partner may be unsubstantial.

One may construct a similar indicator for clients of male sex workers in special surveys of men who have sex with men.

PERCENT OF YOUNG PEOPLE HAVING MULTIPLE PARTNERS IN LAST YEAR

Definition
The percent of young people (15-24) who have had sex with more than one partner in the last 12 months, of all young people surveyed

In a survey among people aged 15-24, respondents are asked about their sexual partnerships in the last year. Those who report more than one partner in the last 12 months constitute the numerator. The denominator is all respondents.

Evaluators should report this result separately for men and women. It may also be constructed separately for those aged 15-19, <15 and 20-24, as appropriate.

This indicator is calculated as:

$$\frac{\text{\# of youth (aged 15-24) who report having sex with more than one partner in the last 12 months}}{\text{Total \# of youth}} \times 100$$

Data Requirement(s)
Self-reported data from survey respondents

Data Source(s)
UNAIDS general population survey; DHS AIDS module; FHI BSS (youth)

Purpose and Issues
Prevention messages for young people tend to begin with abstinence and often focus also on mutual monogamy. But because sexual relationships among young people are frequently unstable, relationships that were intended to be mutually monogamous may break up and be replaced by other relationships in which similar intentions prevail. Particularly in high HIV prevalence epidemics, serial monogamy is not greatly protective against HIV infection. This indicator measures the proportion of young people exposed to more than one partner in the last year, that is, the proportion for whom the "one, mutually faithful partner" message has failed.

This indicator does not distinguish between marital and non-marital partners. It tracks all multiple partnerships, regardless of their relative levels of risk. In the very similar adult sexual behavior indicator (**Percent of Population Who Had Higher Risk Sex in the Last Year**), a distinction is made between marital and cohabiting partners, and all other partner types. This distinction is partly to cope with the measurement challenge posed by men in polygynous societies, who may have multiple partners within marriage. However, because polygyny among men under 25 is extremely rare, the distinction is unnecessary in an indicator for young people.

The indicator also suffers from the expected respondent and social desirability bias. Young people saturated with prevention messages will be highly motivated to underreport partners. Likewise, social pressure for women to give untruthful answers may be strong.

PERCENT OF YOUNG PEOPLE USING A CONDOM AT LAST HIGHER RISK SEX

Definition

The percent of young people (aged 15-24) who have had sex in the last 12 months and used a condom at last sex with a non-marital, non-cohabiting partner, of all young people surveyed

In a general population or targeted youth survey, the interviewer asks all respondents about their sexual partnerships in the last year. For each partner a young person reports, cohabitation status is established. Where a general population survey is undertaken for people aged 15-49, the evaluator can simply stratify the data by age groups to calculate this indicator. The denominator is all young people aged 15-24. The numerator is the percent of those persons using a condom at last sex with a non-marital, non-cohabiting partner in the last 12 months.

This indicator is calculated as:

$$\frac{\text{\# of youth who report using a condom at last sex with a non-marital, non-cohabiting partner}}{\text{Total \# of youth}} \times 100$$

Data Requirement(s)

Self-reported data from survey respondents

Data Source(s)

UNAIDS general population survey; DHS AIDS module; FHI BSS (youth)

Purpose and Issues

The indicator includes the non-marital partners of young people who are currently married, as well as all reported partners of those who are still single or not in a stable enough relationship to be cohabiting with their partner.

The indicator differs from **Percent Using Condoms at Last Higher Risk Sex** in that it includes in the denominator all respondents, rather than just those who report risky sexual activity in the last year.

The evaluator should report the indicator across the 15-24 age range and separately by sex. It may also be reported separately for those aged 15-19, 20-24, and under 15 years, where relevant.

In terms of advocacy, this indicator of young people's sexual behavior can have powerful effects. Where the indicator shows low levels of condom use with higher risk sex among youth, programs will need to focus efforts around abstinence after initiation of sexual activity, but primarily on condom use. But major constraints may exist in programs that wish to avoid addressing youth sexual activity.

Like the previous indicator, **Percent of Young People Having Multiple Partners in Last Year**, this indicator will capture all unmarried people having sex, the proportion of which will generally be fairly high, especially among men. In addition, it will capture married young people having sex outside of marriage.

The indicator suffers from the same reporting bias problems inherent in surveys asking about sexual behavior; depending upon the degree of programs effort saturation and/or existing cultural or religious mores, young people may actually be more willing than adults to report details about their sexual behavior.

PERCENT OF INJECTING DRUG USERS NEVER SHARING EQUIPMENT IN THE LAST MONTH

Definition

The percent of active injecting drug users surveyed who report never sharing injecting equipment during the last month

In a behavioral survey among injecting drug users, respondents are asked about their injecting habits. The indicator excludes those who report sharing needles, syringes, or other injecting equipment at any time in the last month. The denominator is all respondents reporting injecting behavior in the last month.

Questionnaires should specify all the locally relevant types of "equipment" that may result in the exchange of body fluids.

This indicator is calculated as:

# of active injecting drug users who report never sharing injecting equipment during the last month $\times 100$
Total # of respondents who report injecting behavior in the last month

Data Requirement(s)

Self-reported data from survey respondents

Data Source(s)

FHI BSS (injecting drug users)

Purpose and Issues

Sharing injecting equipment between HIV-infected and uninfected drug injectors is an extraordinarily effective way of spreading HIV. Because the risk of contracting infection per single act of risky injection is so high, programs must aim for a complete halt to this behavior, not just for a reduction in the sharing of equipment between drug users.

This indicator measures trends in consistently safe behavior among drug users who continue to inject drugs.

As with all indicators measured among drug injectors, the biggest difficulty is access. Random sampling is all but impossible, and convenience samples are biased in often unpredictable ways. It is therefore difficult to determine the extent to which those surveyed represent the larger population of injecting drug users. Where the representativeness of the sample is variable, trends over time will be hard to interpret.

These surveys yield information from people identified as members of a community of drug injectors. In response to HIV-related interventions, some injectors may possibly stop taking drugs entirely or switch to non-injected drugs. Since the indicator tracks changes in risky injecting practices over time among people who continue to inject drugs, the denominator excludes people who cease to inject.

Some education programs have focused on sterilizing needles between users. Users may continue to inject drugs and even share needles, but may sterilize between uses. However, knowing the success of individual efforts to sterilize equipment is difficult. Experience in some settings has demonstrated that inadequate cleaning of equipment is common, and many programs have ceased to promote equipment cleaning as a prevention method, preferring to concentrate efforts on ending to the sharing of injecting equipment. This indicator includes in its numerator of those with risky behavior all injecting drug users who sterilize, but still share, their equipment.

Because it restricts those included in the indicator to those who have injected in the last month, this indicator is very sensitive to recent trends in injecting practices. Countries with inconsistent policies supporting safe drug injection may see large variations in this indicator. Police crackdowns on users, distributors, or support services such as needle exchange centers may lead to dramatic changes in injecting practices over a very short period of time.

In addition, the indicator is subject to high recall bias. Depending on the local drug scene, drug users may be injecting several times each day. Recalling the circumstances of every act of injection over the past 30 days may be problematic.

PERCENT OF TRANSFUSED BLOOD UNITS SCREENED

Definition

The percent of blood units transfused in the last 12 months that have been adequately screened for HIV according to national or WHO guidelines

The indicator requires three pieces of information: the number of blood units transfused in the previous 12 months, the number of blood units screened for HIV in the previous 12 months, and among the units screened, the number screened up to WHO or national standards.

The number of units transfused and the number screened for HIV should be available from health information systems. Quality of screening may be determined from a special study that re-tests a sample of blood previously screened or from an assessment of the conditions under which screening occurred. Where this approach is not feasible, data on the percent of facilities with good screening and transfusion records and no stock outs of test kits may be used to estimate adequately screened blood for this indicator.

This indicator is calculated as:

$$\frac{\text{\# of blood units transfused in the last 12 months that have been adequately screened for HIV}}{\text{Total \# of blood units transfused}} \times 100$$

Data Requirement(s)

Program records

Data Source(s)

MEASURE *Evaluation* blood safety protocol

Purpose and Issues

Blood safety programs aim to ensure that the overwhelming majority (ideally 100 percent) of blood units are screened for HIV, and that those included in the national blood supply are indeed uninfected. This is de-

monstrably not the case in many countries. Some blood units are not screened at all; others are screened by poorly trained personnel using outdated equipment or insufficient inputs. Moreover, poor blood testing facilities mean that some blood is screened using antibody tests at a time after the donor has become infected with HIV but before the donor has developed antibodies to the virus. Together, these factors mean that a significant proportion of blood units may be classified as safe even though they are infected. This indicator reflects the overall percentage of blood units screened to high enough standards that they can confidently be declared free of HIV.

Where sufficient information exists to construct it, this measure is a strong indicator of the overall safety of the blood supply. However, changes in the indicator could reflect changes in the proportion of blood units screened or changes in the quality of the screening process. The indicator may also reflect a successful campaign to reduce unnecessary transfusions, because the overall number of transfused units would fall and the proportion of those screened to WHO/national standards should rise in consequence. However, the different elements of the indicator should be reported separately for programmatic purposes.

Where health systems are decentralized, or where the private sector is involved in blood screening and blood banking, one may have difficulty obtaining good enough information to construct a robust indicator on a national scale. In this case, selecting sentinel hospitals and laboratories in both the public and the private sector for facility-based surveys of blood transfusion and screening quality will probably be necessary.

PERCENT OF STI PATIENTS APPROPRIATELY DIAGNOSED AND TREATED

Definition

The percent of patients with STIs at selected health care facilities who are appropriately diagnosed and treated according to national guidelines, of all STI patients at those centers

Evaluators collect data through observing and interviewing providers at selected health care facilities offering STI care. Providers are assessed on history taking, examination, and treatment of patients. A provider must score positively on all three items in an interaction with a client for that client to enter the numerator of the indicator.

Protocol researchers have tried several alternative data collection methodologies. Instead of, or in some cases in addition to, observation and provider interviews, data have also been collected through exit interviews with clients and interactions with "mystery" clients – that is, trained assessors posing as clients.

"Appropriate" diagnosis and treatment is assessed according to national guidelines governing STI services. In developing countries, these guidelines will most commonly include protocols for the syndromic management of locally common sexually transmitted pathogens, including treatment with drugs specified in national drug lists. Some countries recognize both syndromic and etiological management as appropriate, according to the diagnostic capacity of the service provider. Where national guidelines are unavailable, WHO guidelines on the syndromic management of STIs may be used to guide assessment of appropriate treatment.

This indicator is calculated as:

$$\frac{\text{\# of patients with STIs who are appropriately diagnosed and treated}}{\text{Total \# of patients with STIs}} \times 100$$

Data Requirement(s)

Assessment of an external expert

Data Sources

WHO/UNAIDS revised guidelines on evaluating STI services; MEASURE Service Provision Assessment (SPA)

Purpose and Issues

STI programs are focusing on syndromic management of STIs as the most practical approach in high prevalence, low resource situations. The shift to syndromic management has increased the potential coverage of care, since such management poses fewer bottlenecks in diagnosis. Training nurses and other health care providers new to the approach and often to STI care in general has required a huge investment.

This indicator reflects the success of that training, combined with efforts to ensure adequate supplies of drugs and other necessary materials to care provision points. It tracks changes in the provision of adequate care to patients seeking care for STIs.

Choosing which STI service delivery points to survey is important. Traditionally, this indicator has applied primarily to public sector STI clinics, because most of the early training in syndromic management was of public sector employees. However, people with STIs often seek treatment in other sectors – either at private sector clinics, from pharmacies or from traditional healers. Some countries have begun to include these sectors in training programs for syndromic management, and evaluations using this indicator have successfully been carried out in these sectors. Service delivery points surveyed should include representative service providers from any sector that has received training in syndromic management of STIs.

This indicator, measured through observation but including provider interviews in the process of data collection for validation purposes, has been widely used and proven feasible. There has been discussion of the

extent to which the direct observation and provider interview methodologies bias data. It is thought that service providers perform better under observation than they normally would, or over-report "correct" diagnosis and treatment, falsely diminishing the gap between knowledge and practice. However, client exit interviews and mystery patient methodologies, as well as proving feasible, have demonstrated that the biases are smaller than was assumed. The gap between knowledge and practice in the area of treatment often shaped results because service providers do not follow "correct" protocols simply because they know drugs are unavailable or unaffordable. We thus recommend that evaluators present this indicator with indicators of drug availability.

As with all composite indicators, improvements in some areas may mask deterioration in others. If service in one area is poor, the facility will score poorly on the indicator, even if service provision in other areas has progressed significantly. Program managers need scores reported separately by area of knowledge and performance so that they may identify areas of weakness and may improve program performance.

PERCENT OF STI PATIENTS RECEIVING ADVICE ON CONDOM USE AND PARTNER NOTIFICATION AND REFERRAL TO HIV TESTING SERVICES

Definition

The percent of patients with STIs who are given advice on condom use and partner notification and referred for HIV testing

Previous indicators only included the first two elements of this indicator. A health care provider must score positively on both condom advice and partner notification advice for the client to enter the numerator for this indicator. The current indicator includes a third element: referral for voluntary testing for HIV. However, if national policy does not include referring STI patients for HIV counseling and testing, or if VCT services are unavailable and not actively promoted by national AIDS and STI programs, the indicator should exclude referral for counseling and voluntary HIV testing. Health facility surveys through direct observation of interaction between care providers and clients yield data for this indicator.

This indicator is calculated as:

$$\frac{\begin{array}{l}\text{\# of patients with STIs given advice} \\ \text{on condom use and partner notification} \\ \text{and referred for HIV testing}\end{array}}{\text{Total \# of patients with STIs}} \times 100$$

The evaluator should report different components of this indicator separately, for reasons given below.

Data Requirement(s)

Assessment by an external expert

Data Sources

WHO/UNAIDS revised guidelines on evaluating STI services; MEASURE Service Provision Assessment (SPA)

Purpose and Issues

By promoting condom use and by encouraging the treatment of partners to avoid reinfection, STI services seek to prevent the recurrence of STIs, not just to treat them. Increasingly, STI care serves as an entry point for referral for voluntary testing for HIV. This indicator measures the extent to which these aspects of STI service provision are functioning.

If a client is at an STI clinic, previous efforts to promote safe behavior have failed him or her. This measure does not evaluate the success of prevention initiatives, merely the extent to which service providers are complying with standards.

The extent to which the direct observation methodology biases data has caused concern because researchers assume that service providers perform better under observation than they normally would. Also, it is suggested that exit interviews with clients may be a more cost-effective method than observed interactions in compiling this indicator. However, clients may misreport the actual content of counseling. Further research is needed to determine the reliability of exit interviews in collecting data for this indicator.

Condom promotion, advice on partner referral, and referral for HIV testing are in fact quite distinct activities. The value of an aggregate indicator in this field is therefore somewhat limited, at least to program staff. In addition, referral to HIV testing services will depend upon the availability of those services locally. And the addition of this component will disrupt trends over time in those countries where a different indicator has been calculated in the past. For these reasons, the evaluator must take special care to report separately the three elements of this indicator.

PERCENT OF HEALTH FACILITIES PROVIDING STI SERVICES WITH ADEQUATE DRUG SUPPLY

Definition

This indicator measures the percent of health facilities providing STI care that have a current supply of essential STI drugs and report no stockouts lasting longer than one week in the preceding 12 months

Countries promoting syndromic management of STIs usually have protocols for the prescriptions of drugs by syndrome as well as a national essential drug list that includes the relevant drugs. Drugs necessary to treat each of the important syndromes should appear in the stock-check for this indicator. A survey of randomly selected facilities providing STI services checks for current supplies of designated drugs. Clinic management is questioned about stockouts in the last 12 months, and clinic stock records are reviewed for that period. Client numbers are also recorded. The sampling frame for the selection of sites may include private clinics and hospitals and non-government services, as well as public facilities.

Evaluators may construct the indicator weighting each facility by its client load because a rupture of stock at a small rural clinic will have less impact on the epidemic at a national level than a stockout in a large urban clinic that sees many times more patients.

This indicator is calculated as:

$$
\frac{\text{\# of health facilities providing STI services that have a current supply of essential STI drugs and report no stockouts}}{\text{Total \# of health facilities providing STI services}} \times 100
$$

Depending on national policy, the indicator may include a variety of outlets providing services for HIV care, such as integrated reproductive health services, private sector facilities, and pharmacies with special training in STI care provision.

Data Requirement(s)

Assessment by an external evaluator

Data Sources

WHO/UNAIDS revised guidelines on evaluating STI services; MEASURE Service Provision Assessment (SPA)

Purpose and Issues

Correct history-taking, diagnosis, and prescription are all very well, but if drugs are unavailable, these procedures will not translate into cases cured and will therefore not reduce the likelihood of HIV infection.

National AIDS programs engaged in improving STI services have put time and money into improving drug distribution services and in attempting to ensure adequate manufacturing or importing of drugs for the syndromic management of STIs. This indicator measures the effectiveness of those efforts in ensuring that service providers are consistently supplied with the drugs they need to work efficiently.

This indicator is a good measure of consistent supplies of drugs to STI service facilities; it provides a minimum measure of the availability of drugs. However, clients very often buy drugs from other sources, even when they have been to an STI facility for diagnosis. Indeed, in countries where control of drug supplies are lax, a stockout in a public clinic may simply mean that the supply of drugs has been diverted to another nearby outlet. This laxity will likely affect the cost of the drug to the client (and therefore accessibility), but it may not affect the physical availability of the drug.

Again, the selection of STI facilities may have a major influence on the indicator. The facility survey should attempt to include a mix of all major provider categories in both the public and the private sectors.

NUMBER/PERCENT OF HEALTH FACILITIES WITH THE CAPACITY TO DELIVER APPROPRIATE CARE TO HIV-INFECTED PATIENTS

Definition

The number and percent of health care facilities at different levels of the health care system that have the capacity to deliver appropriate palliative care, treatment for opportunistic infections, and referral for HIV-infected patients, according to national guidelines

A health facility survey that includes facility inspection, interviews with service providers, and records reviews assesses health facilities against a standard checklist. The checklist, which will be modified according to local standards, will differ according to the level of the institution within the health care system. It will typically include the availability of trained staff, the adequacy of diagnostic facilities, the adequacy of sanitation, the adequacy of nursing care, procedures for record keeping, preventative counseling, and referral to higher level care and community support organizations as appropriate.

The assessment of "adequate" or "appropriate" conditions and services should follow national guidelines for care of HIV-infected patients. The absence of such guidelines in itself indicates that care and support services for HIV-infected people are likely to be inadequate. However, where they do not exist, one may substitute international standards currently being developed by WHO to determine standards against which to measure facilities.

This indicator excludes the availability of drugs and procedures to prevent accidental transmission of HIV within the health care setting because separate indicators cover this availability.

The indicator is the number of health facilities matching or exceeding the minimum score for adequate capacity to manage HIV-infected patients, divided by the total number of health facilities surveyed. For program purposes, it should be disaggregated by level of health facility as well as by area of service provision.

This indicator is calculated as:

$$\frac{\text{\# of health care facilities with the capacity to deliver appropriate palliative care, treatment for opportunistic infections, and referral for HIV-infected patients}}{\text{Total \# of health care facilities}} \times 100$$

Data Requirement(s)

Assessment by external evaluator of adequacy of care to HIV infected patients

Data Sources

WHO draft protocol for the evaluation of HIV/AIDS care and support; UNAIDS protocol for evaluation of care and support

Purpose and Issues

In the early years of the HIV epidemic, a high proportion of patients with HIV-associated conditions were automatically referred to tertiary level institutions because health services at other levels had neither the trained personnel nor the capacity to cope with them appropriately. Even guidelines on what constituted "appropriate" treatment were rarely available. The constant referral to higher levels of care clearly led to inefficient use of resources within the health system.

In recent years, attempts have been made to ensure that HIV-related conditions are dealt with at appropriate levels within the health system, with referrals in both directions when necessary. Many countries have produced national guidelines to help guide service providers in the appropriate care of HIV-infected patients. Palliative care and treatment for common and minor opportunistic infections may be given at the primary level, while more complex opportunistic infections may be referred to higher levels of the health care system. Referrals should also be made for social and psychological support where appropriate.

This indicator measures the extent to which health services have the capacity to meet treatment, care, and referral needs of HIV-infected patients at appropriate levels of the health care system, according to national guidelines.

This indicator is a compendium of many different aspects of care and service provision, all of which must score a minimum amount if the indicator is to include the facility in its numerator. Because services tend to improve unevenly, especially in resource constrained settings, the resulting indicator may remain low for some time. Disaggregation of the indicator will indicate the areas in which services have improved and those in which they continue to lag.

The scoring of the components of the indicator will necessarily include a measure of subjectivity. This subjectivity may influence comparisons between different countries, as well as trends over time if the monitoring team changes.

Because it includes facilities at different levels of service provision, the indicator is not weighted by client load. Weighting by client load is likely to give tertiary institutions and reference hospitals excessive influence in the indicator, despite the fact that most patients first come into contact with the health system at the primary level.

HIV Prevalence among Pregnant Women 15-24 Years Old

Definition

The percent of blood samples taken from women aged 15-24 who test positive for HIV during routine sentinel surveillance at selected antenatal clinics

The data for this indicator are obtained from the national sentinel surveillance system for HIV, and the indicator is calculated through unlinked anonymous testing for HIV of blood samples taken from women at sentinel antenatal clinics chosen to reflect urban, rural, ethnic, and other socio-geographic divisions in a country.

Even where programs exist that simultaneously offer counseling and voluntary HIV testing for pregnant women to reduce mother to child transmission, only the results of unlinked, anonymous screening of blood taken for other purposes should be used in calculating this indicator of HIV prevalence. Refusal and other participation bias are considerably reduced in unlinked anonymous HIV testing compared with other forms of testing.

This indicator is calculated as:

$$\frac{\text{\# of pregnant women aged 15-24 who test positive for HIV}}{\text{Total \# of women aged 15-24 tested}} \times 100$$

Data Requirement(s)

Results of test for HIV sero-positivity

Data Sources

UNAIDS/WHO Second Generation Surveillance; WHO guidelines for HIV surveillance

Purpose and Issues

Women who are pregnant have by definition had unprotected sex sometime in the last ten months. Levels of HIV infection in these women do not reflect levels among women who are not having sex, among women who are infertile, or among women who are systemati-

cally using contraception, including barrier methods such as condoms which also prevent HIV transmission.

Confining the indicator to women aged under 25 aims to give a picture of recent trends in infection. Most infections in this age group are relatively new, and data from these younger women are also less subject to bias than data for the whole reproductive age span. The indicator is reported for women aged 15-24. However, we strongly recommend that the evaluator report two separate figures: one for women aged 15-24 and one for women across the whole reproductive age range of 15-49. Because many countries have in the past failed to report HIV prevalence broken down by age, it is important to continue to report a figure for HIV prevalence across 15- to 49-year-olds, to allow for the comparison of trends over time.

Additional information may be gained by looking at HIV prevalence by parity of mother. Such information is often routinely collected in sentinel surveillance, and analysis of trends among women of parity 0 and 1 combined is a good additional indicator of trends in HIV incidence among young women.

The indicator gives a fairly good idea of relatively recent trends in HIV infection nation-wide in countries where the epidemic is heterosexually driven. It is less reliable as an indicator of overall epidemic trends in areas where the bulk of HIV infection remains confined to sub-populations with especially high-risk behaviors. Even countries with generalized heterosexual epidemics have wide regional, ethnic, or other differences in trends in HIV infection. These differences will be lost when data are aggregated into a single national figure. For program purposes, prevalence should thus always be reported separately by site as well as by a single national figure. Evaluators should take care in reporting HIV prevalence estimates by sites, however, given the possible political sensitivity of results.

In the past, sample sizes in regular sentinel surveillance have been selected to measure changing trends across

the whole age range of 15-49. Numbers in each five-year age band may have been too small to yield any reliable trend data, particularly at individual sentinel sites. To construct a reliable indicator around the narrower age range, larger sample sizes in the younger age groups will be needed.

Clearly, trends in HIV infection among pregnant women will not adequately reflect some of the most important changes in behavior supported by AIDS prevention programs – abstinence and consistent condom use in all populations and not simply the antenatal care clients. Trends in HIV infection are beset by a number of biases, as described above. Prevalence among pregnant women reflects trends in prevalence in the general population, but does not an accurately reflect overall levels in all women, let alone in all men.

Evaluators should thus report prevalence data together with behavioral data (such as mean age at first sex or condom use at last sex) for better explanatory power.

HIV PREVALENCE IN SUB-POPULATIONS WITH HIGH-RISK BEHAVIOR

Definition

The HIV prevalence among members of a defined sub-population at higher risk of contracting or spreading HIV.

Tracking HIV in sub-populations can be logistically and ethically difficult, especially if the groups are marginalized or their activities are illegal. Sampling and estimation of total population sizes are key issues. An understanding of how the sampled population relates to any larger population sharing similar risk behaviors is critical to the interpretation of the indicator. For some groups, population-based sampling strategies will be necessary. In other cases, sentinel sites are available. Sentinel sites for these populations tend to be linked to the provision of health services, for example, a men's health clinic in an area with a high concentration of gay sex bars, or a drug rehabilitation center.

The indicator is the number of members of the at-risk sub-population testing positive for HIV at sub-population sentinel sites, divided by the total number of members of the at-risk sub-population tested for HIV.

This indicator is calculated as:

$$\frac{\text{\# of members of at-risk sub-populations testing positive for HIV}}{\text{Total \# of members of at-risk sub-population tested}} \times 100$$

Data Requirement(s)

Results of test for HIV sero-positivity

Data Sources

UNAIDS/WHO Second Generation Surveillance guidelines; FHI guidelines on sampling in sub-populations

Purpose and Issues

In countries with concentrated epidemics, tracking of HIV infection among pregnant women may be a waste of resources. In any case, the bulk of interventions in concentrated epidemics often focus on the behaviors or groups contributing most to the expansion of the epidemic. In a concentrated epidemic, these generally include one or more of the following: injecting drug users, men who have sex with other men, sex workers, and frequent clients of sex workers.

The design of a second generation surveillance system should take into account the epidemic state. In countries with low-grade or concentrated epidemics, surveillance for the HIV virus as well as behavioral surveillance should focus on those groups where both infection and interventions are concentrated. Changes in HIV prevalence in these groups should reflect the success or failure of prevention attempts.

Because of the difficulties in access to sub-populations, the biases in sub-population sero-surveillance data are likely to be far greater (and much less predictable) than in data from a more generalized population such as women at antenatal clinics. Where sentinel sites provide health services to the sub-population in question, for example, the use of the facility may be associated with problems that are themselves related to HIV infection.

Minimizing biases associated with age is especially difficult, because the age of participation in especially high-risk behaviors may vary widely. It is therefore, undesirable simply to restrict the analysis to young people as it is in ANC sentinel sites.

Despite these difficulties, tracking HIV infection in those with higher risk behaviors in concentrated epidemics is essential. The information will not be perfect, but some measure of progress or lack thereof will be essential to maintain support for prevention programs in critical sub-populations.

PERCENT OF CHILDREN UNDER 15 WHO ARE ORPHANS

Definition
The percent of children under 15 whose mother, father, or both parents have died.

In a household survey or a national census, respondents are asked the ages of all children in the household and whether the mothers and fathers of those children are alive. Those children who are currently under the age of 15 and whose mother or father or both are dead form the numerator for this indicator. The denominator is all children currently under 15 listed by respondents in the survey.

Breaking the results down into maternal orphanhood, paternal orphanhood, and double orphanhood is useful.

This indicator is calculated as:

$$\frac{\text{\# of children under 15 whose mother, father, or both parents have died}}{\text{Total \# of children under 15 listed by respondents}} \times 100$$

Data Requirement(s)
Responses on household surveys

Data Sources
Household schedule in UNAIDS general population survey; DHS household schedule; Census data

Purpose and Issues
HIV is changing the face of adult mortality in many communities, killing men and women at just the ages when they normally form families and bring up children. Their deaths leave behind orphans who must be cared for, generally by other members of the community. The social and economic impact of rising orphanhood can be considerable; national AIDS programs tracking orphanhood will be better equipped to plan for impact mitigation efforts. This indicator tracks levels of orphanhood in a country.

Data on an increase in orphanhood can be a very powerful indicator of the impact of an AIDS epidemic. Besides tracking the impact of AIDS deaths on communities, this indicator also has multiple advocacy uses.

One limitation of this measure is that it cannot distinguish AIDS-related orphanhood from orphanhood due to other causes. However, because young adult death was stable or falling in most countries for some years before the arrival of HIV, we can reasonably assume that the bulk of any rise in orphanhood over baseline levels is attributable to HIV.

Orphans may be more mobile than other children. Those most in need of care may be in child-headed households that do not qualify for inclusion in a household survey. Street children living in orphanages will also be missed. Households with AIDS-related deaths often completely disintegrate following the death of heads, and children are sent to live with relatives in the same or another area. Using a household survey and asking about whether the parents are still alive will help alleviate the primary household disintegration issue.

Definitions of orphanhood differ among countries. For example, in some countries, the legal definition includes all children under 18 who have lost either or both parents, whereas in others it includes all children under 15 who have lost their mother. We suggest that the standard definition given in this indicator allow for comparison across populations. However, countries may also wish to compile an indicator based on their own national definition of orphanhood. The methodology for constructing the indicator remains unchanged.

Part III.D
Safe Motherhood

- Existence of a safe motherhood strategic or operational plan to promote access and/or quality of safe motherhood services
- Maternal Neonatal Program Index (MNPI)
- Number of facilities per 500,000 providing essential obstetric functions
- Percent of facilities that conduct case review/audits into maternal death/ near miss
- Percent of pregnant women attending antenatal clinics screened for syphilis
- Percent of women with obstetrical complications treated within two hours at a health facility
- Cesarean sections as a percent of all live births
- Case fatality rate (CFR) – all complications
- Percent of audience that know three primary warning/danger signs of obstetric complications
- Percent of women attended at least once during pregnancy for reasons related to the pregnancy
- Percent of women who were given or purchased malaria prophylaxis/treatment during their most recent pregnancy
- Percent of pregnant women who receive antihelminthic treatment during pregnancy
- Percent of births attended by skilled health personnel
- Percent of women attended during the postpartum period by skilled personnel
- Maternal mortality ratio (MMR)
- Met need for essential obstetric care (EOC)

The inauguration of the Safe Motherhood Initiative in Kenya in 1987 marked the beginning of concerted international efforts to reduce maternal mortality. Since that time, reducing maternal mortality has continued to be the goal of many international health programs. International efforts aim to achieve a 75 percent reduction in maternal mortality ratios between 1990 and 2015 (WHO, 2001c).

Although safer motherhood has remained high on the political agenda, the scope of what constitutes safer motherhood has changed considerably over the past 13 years. A major factor has been the incorporation of a human rights approach into the definition of Safe Motherhood following the agenda set at the International Conference on Population and Development (ICPD). By defining maternal death as social injustice, programs for "Safer Motherhood" are able to invoke a much broader range of political, social, and economic initiatives than was previously possible (UNFPA et al., 1997).

Policies and strategies to achieve safe motherhood have also changed as knowledge and understanding about the determinants of maternal health have become clearer. Major new policy initiatives based on the lessons learned after the first decade were condensed into a series of ten action messages or "Priorities for the Next Decade" by an international panel of experts in Sri Lanka in 1997 (UNFPA et al., 1997). Two of these lessons warrant particular mention because of their subsequent role in redefining safe motherhood policy and strategy.

The first of these – that the difficulty of predicting complications means that "every pregnancy faces a risk" – has shifted the emphasis from targeting women "at risk" towards providing universal access to care. With the understanding that maternal survival is an issue of access to health infrastructure and timely emergency care, current program efforts have redirected towards providing all women with a basic "minimum package" of services. At the first-referral level, basic essential obstetric care (BEOC), should include services for a normal pregnancy as well as the early detection and management of complications. The provision of surgery and blood transfusion should be made available at the secondary level in comprehensive essential obstetric care (CEOC) facilities (UNICEF, WHO, and UNFPA, 1997).

The second lesson – that the strategy of training traditional birth attendants (TBAs) with limited midwifery skills has not reduced maternal mortality – underlies the current global initiative to increase the proportion of deliveries that take place with a skilled birth attendant (Berer and Ravindran, 2001; Graham and Bell, 2000).[1] The assumption behind this approach is that a skilled attendant (a doctor, midwife, or nurse with midwifery skills) with a higher level of training and expertise should be able to manage and refer pregnancy complications more effectively and hence should reduce maternal deaths. Because the majority of both maternal and newborn deaths occur between the onset of labor and the early postnatal period, ensuring the presence of a skilled attendant at birth is widely seen as the single most critical intervention for reducing both maternal and newborn mortality (Graham, 2001).

Over the past 13 years, many other new priorities and challenges have also emerged that need to be addressed under the safe motherhood umbrella. Foremost among these is the HIV/AIDS pandemic, which in some African countries affects up to 30 percent of antenatal patients and which urgently require major new interventions (Berer and Ravindran, 2001). Another major development is the increasing focus on newborn care. Awareness that relatively simple interventions both before and after birth may substantially reduce the estimated eight million perinatal and neonatal deaths has given rise to many more programs that target improving newborn survival. For further discussion of newborn care, see Part III.E.

---

[1] The term "skilled attendant" refers exclusively to people with midwifery skills (for example, doctors, midwives, nurses) who have been trained to proficiency in the skills necessary to manage normal deliveries and to diagnose or refer obstetric complications (WHO, 1999b.).

In 1992, WHO, UNICEF, UNFPA, and the MotherCare Project began to advocate using program-level indicators[2] to measure aspects of facility care, such as the quality of services, women's access to services, and women's utilization of those services. Program-level indicators have several advantages over population-based measures. They can be used for short term monitoring; they provide suitable information for local management decision making; and they can be derived from existing data at low cost. Several indicators suggested by these groups are included in this manual. Until recently, experience with using these indicators has been limited, but many countries are now currently carrying out needs assessments and program monitoring using these indicators. The number of publications describing the strength and limitations of the indicators is growing (Bailey et al., 2001; Pathak et al., 2001; Ronsmans et al., 1999)

Conceptual Model

Several different models or frameworks exist to help program managers and communities understand the determinants of maternal mortality (Campbell et al., 1997; McCarthy and Maine, 1992; Thaddeus and Maine, 1994; Koblinsky et al., 2000). The "Three Delays Model" identifies the points at which delays can occur in the management of obstetric complications at the community and facility level.

The first "delay" (the delay in deciding to seek care) may relate to a number of factors, including the lack of knowledge about obstetric danger signs, community perception of poor quality facility care, or the lack of health services availability which increases the opportunity costs and therefore reduces the likelihood of care seeking.

The second "delay" (the delay in identifying and reaching a medical facility) relates to the geographical proximity and accessibility of health services, and includes factors such as the availability of transportation.

The third "delay" (delay in receiving appropriate care at health facilities) is related to factors in the health facility, including the availability of staff, equipment, and resources as well as the quality and (in some cases) the cost of services.

Indicators for measuring aspects of each of these delays can be found in the following section as well as in other parts of the *Compendium*.

---

[2] Examples include met need and unmet need for obstetric care, the cesarean section rate, percentage of women with a trained assistant at delivery, and place of delivery (as measures of women's access to and utilization of services), and the case fatality rate and referral rate (for quality of care).

Figure III.D.1 The Three Delays Model
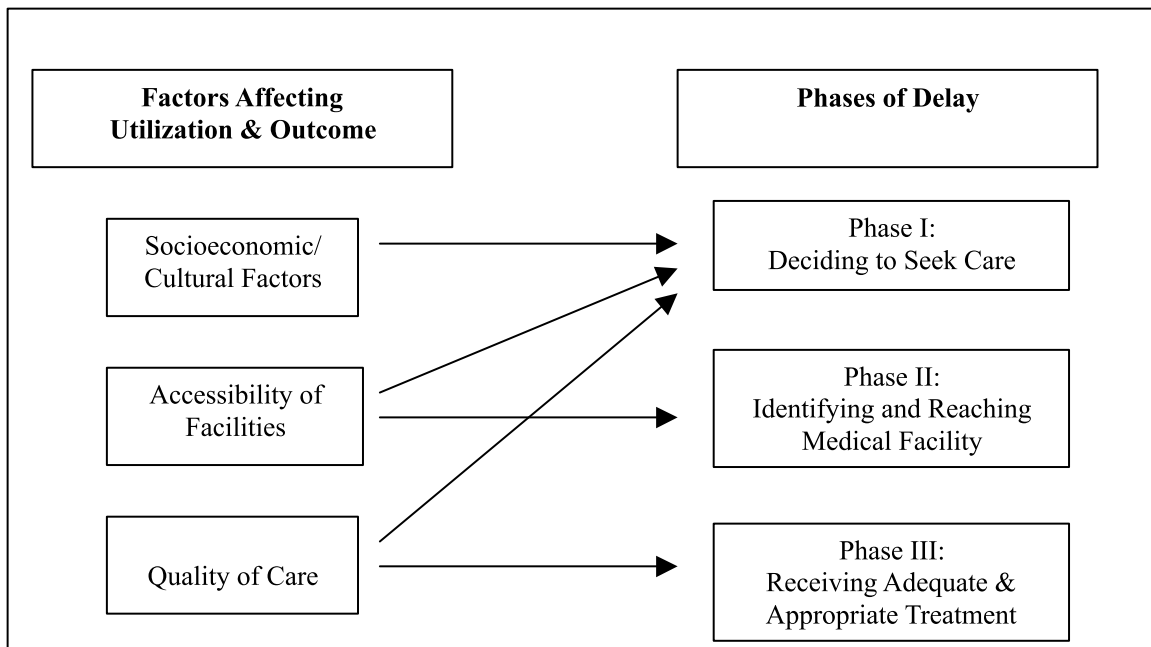
Source: Thaddeus and Maine, 1994

Methodological Challenges of Evaluating Maternal Health Interventions

In addition to the changes in the definition, policies, and strategies as well as the emergence of new public health problems that drive the need for an increasingly wide range of indicators, monitoring and evaluating safe motherhood programs pose a number of inherent methodological challenges. These include, but are not limited to, the following issues.

- **Maternal mortality is difficult to measure, and estimates of maternal mortality should not be used for monitoring purposes.**

Maternal mortality estimates are valuable for advocacy purposes because they directly measure progress towards the goal of reducing maternal mortality. Mortality estimates do, however, have a number of inherent methodological weaknesses that limit their use for monitoring purposes; they are costly, they do not explain the causes of maternal deaths, and they cannot detect short-term change (Graham, 2001).

Few developing countries have registration systems with sufficiently wide coverage to provide accurate national estimates of maternal mortality. Alternative approaches to deriving estimates, such as surveys and the sisterhood method, also have limitations in that the estimates are relatively imprecise and relate to periods several years before the survey. Even in countries where the maternal mortality is high, maternal deaths are rare events; therefore, surveys are very costly because of the need for large sample sizes to provide a statistically reliable estimate. The wide confidence limits on the estimate also make it very difficult, if not impossible, to assess whether change has occurred over time. For these reasons, maternal mortality estimates, if required, should be measured only infrequently (e.g., once a decade), and program-level indicators that measure the availability, use, and quality of care are recommended for monitoring purposes (AbouZahr, 1999).

- **Maternal morbidity is difficult to define, interpret, and measure.**

Maternal morbidity is much more common than maternal death; thus, the prevalence of maternal morbidity provides a conceptually appealing alternative outcome to measure. Moreover, relatively little is known about the burden of reproductive morbidity; more work is needed to explore the dimensions and determinants of

the problem as well as to evaluate the effectiveness of interventions (UNFPA et al., 1997).

The link between morbidity and mortality is not straightforward. Safe motherhood interventions primarily offer secondary prevention; that is, they prevent deaths from complications rather than preventing the complications themselves. Furthermore, unlike death, which has a very defined outcome, measures of morbidity are difficult to define and thus to measure. Even persons with medical training may misclassify complications; consequently, generating any meaningful comparative measures is difficult (Fortney and Smith, 1999).

- **Outcomes need to be measured for two individuals, the mother and baby.**

Safe motherhood programs need to consider the outcomes for two individuals: the mother and baby. Under most circumstances interventions that benefit or harm the mother similarly affect the baby and vice versa. Some exceptions are notable. For example, a cesarean section for fetal distress may be critical to ensure a good neonatal outcome but may more negatively influence the mother's health than a normal vaginal delivery will.

- **The provision of appropriate maternity care is a complex process that requires multiple indicators to monitor.**

Unlike most areas of public health, providing appropriate maternity care is a complex process that involves a wide range of preventive, curative, and emergency services as well as several different levels of care (from the community to the facility and beyond). The occurrence of an emergency sets into motion a complex chain of events to ensure that a woman receives adequate care. First, the family needs to recognize the problem and be able to access the appropriate services. Second, the equipment, supplies and medicines must be available at the facility to enable the care provider to make the correct diagnosis and to provide appropriate treatment promptly. If definitive care cannot be provided at the first level, then transport needs to be available quickly to take the woman to a higher level of care that must also deliver the appropriate services. Problems at any one of these stages may mean that the woman receives substandard care, which may be of critical importance in determining the outcome. From a program perspective, a series of indicators is required to reveal whether a problem occurs on the "demand" or "supply" side of

the equation, and hence, whether the interventions need to address community mobilization, behavior change, health system performance, or a combination of these factors.

- **Interpreting whether outcomes are attributable to program interventions is difficult, because most interventions consist of "bundled" services.**

Demonstrating change as a result of a safe motherhood program is difficult because programs usually provide a package of care to communities rather than providing one single intervention. Therefore, such programs do not lend themselves easily to two common experimental designs: randomized control trials or cluster randomized community-based trials. Many programs adopt "before-after" designs for evaluation purposes that can demonstrate "plausible association" but that fall short of determining causality (UNFPA et al., 1997).

The Selection of Indicators

The indicators in this section of the *Compendium* are intended mainly for use at the national level or in the context of large-scale programs. However, many can serve in a much wider monitoring and evaluation context. A small expert group currently working in monitoring and evaluation of safe motherhood/newborn health programs selected these indicators in wide consultation with other experts working in the field of maternal health on the basis that they:

- Are widely used by international organizations or ministries of health;
- Have a relatively strong link to health or mortality outcomes; and
- Will likely provide valid comparisons at a national and international level.

Not all indicators included in this section are equally strong or provide the same quality of information. Certain indicators (for example, **Percent of Pregnant Women Attending Antenatal Clinics Screened for Syphilis**) are included because of the potential importance of the information, even though the feasibility of collecting valid information at a national level may be low. At least one indicator is included to represent each element of the proposed safe motherhood framework (see Figure III.D.1), and under some headings, no "strong" indicators were available. In addition, the program-level indicators are included because they are now

used rather widely, and they provide useful information about local planning and decision-making. However, several are clearly not intended for national level use.

The indicators included in this section focus on a rather narrow definition of safe motherhood; they do not address related issues of HIV/AIDS, STIs, or nutritional status, even though these factors clearly have a profound impact on maternal and newborn health outcomes. Priority is given to those indicators currently in use to monitor safe motherhood programs and those most closely related to maternal outcome. Although we recognize the close links between safe motherhood and these other areas of reproductive health, indicators in these related areas appear in their respective sections.

Ideally, those working in safe motherhood will eventually arrive at a consensus as to which indicators should serve for national monitoring of safe motherhood programs, as has been achieved for monitoring national AIDS programs (UNAIDS, 2000). The indicators in this section represent one step toward such consensus. However, programs need to develop their own set of indicators according to the objectives of the program and the interventions designed to achieve those objectives.

EXISTENCE OF A SAFE MOTHERHOOD STRATEGIC OR OPERATIONAL PLAN TO PROMOTE ACCESS AND/OR QUALITY OF SAFE MOTHERHOOD SERVICES

Definition

The degree of explicit support for access to and/or quality of safe motherhood programs on the part of the government and other bodies, including service delivery organizations

Most, but not all, developing countries now have some national FP/RH law, policy, or strategy in place. Safe motherhood policies and plans may be separate from or included within the larger RH policies or strategic plans.

Data Requirements

This qualitative indicator is based on the existence of a safe motherhood plan. Evaluators assign a "yes" value if a strategic or implementation plan exists. Sometimes safe motherhood strategies are incorporated into reproductive health or maternal and child health implementation plans; thus, evaluators should assess these plans to determine if the objectives and corresponding strategies adequately address safe motherhood.

Evidence of an approved plan for safe motherhood with evidence for approval (or submission for approval). In addition, supporting documentation should include the plan, where and by whom it was issued or published, and how the plan promotes access and/or improves the quality of safe motherhood services.

Data Source(s)

Documents from the government organization designated as responsible for coordinating safe motherhood or reproductive health. Content analysis of the plan document should determine whether the plan: (1) defines the objectives of the country's safe motherhood program; (2) defines a clear strategy for attaining these objectives; (3) establishes an organizational structure for the program which is consistent with the strategy and which covers both public and private sectors, including women's groups; and (4) estimates and projects the resources required to implement the strategy, and specifies how these resources are to be secured.

Purpose and Issues

This indicator measures the degree of explicit support for access to and/or quality of safe motherhood programs on the part of the government and other bodies, including service delivery organizations. This indicator tells us if policy is translated into a strategic or implementation plan. Its purpose is to measure whether the safe motherhood or pregnancy program has developed a clear view of its mission and objectives and the strategies for attaining them. Strategic implementation planning at the national level requires the participation of various government ministries or departments, including the health, finance, planning, information, education, interior ministries, as well as important private groups (NGOs and commercial establishments), women's groups, and religious and civic organizations.

Maternal Neonatal Program Index (MNPI)

Definition

This indicator is a score (ranging from 0-100) that measures the strength of the national maternal and neonatal health program of a given country based on five main areas: policy and support services, facility capacity, access to services, care received, and family planning. The five main areas cover 13 components:

- Capacities of health centers;
- Capacities of district hospitals;
- Percent of women with access;
- Care at antenatal visits;
- Care at delivery;
- Care for newborns;
- Family planning at health centers;
- Family planning at district hospitals;
- Policies toward safe pregnancy;
- Resources;
- Information, education;
- Training; and
- Monitoring and evaluation.

Within each component, the researcher averages items to produce a component score, and converts these scores to a 0-100 scale. The index also yields a total score, which is simply the mean for the 13 components with equal weight for each component (Bulatao and Ross, 2000).

Data Requirements

Responses to a detailed questionnaire composed of 81 items from selected key informants (experts from Ministries of Health, medical schools and universities, non-governmental and community organizations, and donors). Besides rating current program adequacy, experts rate each item on the questionnaire as of three years prior to the survey.

Data Source(s)

The MNPI questionnaire completed by 10-25 individuals per country.

Purpose and Issues

The purpose of the MNPI is to assess the strength of a national maternal and neonatal program and measure changes over time. More specifically, the MNPI is intended to measure the effort put into reducing the maternal/neonatal mortality and morbidity in a given country. The index is designed to assess only the program inputs, processes, and outputs as they relate to the "supply" or program side of the conceptual framework (Ross, Campbell and Bulatao, 1999). The MNPI also provides a measure by which to make cross-country and regional comparisons. The MNPI is not designed to provide a single measure of the quality of maternal care; rather, it provides many measures that collectively define a broad standard programs should meet (Bulatao and Ross, 2000).

The MNPI instrument relies on expert judgments, replicating an approach used in family planning and HIV/AIDS. (See **Family Planning Program Effort Index** and **AIDS Program Effort Index (API)** in Parts III.B and III.C, respectively.) Standards, however, are ultimately subjective, resting on the knowledge and expertise of the raters, who are different in each country. Whereas the data collection protocol calls for using raters with varying backgrounds (at least ten per country), validating their ratings is difficult (Ross, Campbell and Bulatao, 1999).

Another limitation is the difficulty of correlating the index closely to a reduction in maternal mortality, largely because maternal mortality levels are hard to determine accurately (Bulatao and Ross, 2001).

NUMBER OF FACILITIES PER 500,000 PROVIDING ESSENTIAL OBSTETRIC FUNCTIONS[3]

Definition

The number of facilities providing essential or emergency obstetric signal functions per 500,000 population

Essential obstetric signal functions are defined as:

- Administration of parenteral antibiotics;
- Administration of parenteral oxytocic drugs;
- Administration of parenteral anticonvulsants for pregnancy induced hypertension;
- Performance of manual removal of placenta;
- Performance of removal of retained products (e.g., vacuum aspiration);
- Performance of assisted vaginal delivery (e.g., vacuum extraction, forceps);
- Performance of surgery (e.g., Cesarean section); and
- Performance of blood transfusion.

Facilities are divided into those that provide "basic" essential obstetric care (EOC) and "comprehensive" EOC. If a facility has performed each of the first six functions *in the past three months*, it qualifies as providing basic EOC. If it has provided all eight of the functions, it qualifies as a "comprehensive" EOC facility.

Data Requirements

Count of the facilities meeting the requirements for "basic" and "comprehensive" EOC

Data Sources

Facility surveys that examine medical records or service statistics. Ideally, records should provide the essential obstetric signal functions. Personal interviews with knowledgeable staff who attend obstetric patients are a second, albeit, potentially more biased source of information than written records are.

Purpose and Issues

This indicator demonstrates the existence of life-saving obstetric care services. It distinguishes between "ba-

sic" and "comprehensive" care services to emphasize that maternal lives can be saved not only in hospitals providing all the services listed above, but also at health centers or smaller hospitals that do not.

The list is intentionally brief to facilitate assessment and monitoring; it does not constitute the complete list of services that either a basic or comprehensive EOC facility should provide. Valuable services are omitted in the definition of an EOC facility. For example, use of anesthesia is not included, although assumed necessary for obstetric surgery.

To qualify as providing quality services, a facility must have a safe and secure blood bank with universally accepted screening tests and a supply of blood. Ideally, the facility has the signal functions available 24 hours a day, 7 days a week, at least in the comprehensive facilities.

The causal link between maternal deaths and this indicator, although logical, has not been demonstrated. Clearly, EOC services must exist to save many women's lives.

This indicator is relatively easy to produce, but it should reflect how facilities are actually functioning and not how they are supposed to function. For example, providers may lack confidence in their skills and refer patients to a higher level, although they are otherwise equipped to treat these patients.

Generally, facility-based assessments cover all the facilities in a specific area. Private facilities may be more reluctant to collaborate than may public facilities. Also, samples of facilities generalizable to a national level such as the Service Provision Assessment (SPA) are possible, but may not always include all the signal functions listed above (MEASURE DHS+, 2000).

---

[3] Much of the text for this indicator comes from Maine, McCarthy, and Ward, 1992 and UNICEF, WHO, and UNFPA, 1997.

This indicator should respond to changes within a fairly short period of time (e.g., 6-12 months).

Generally, this indicator applies to a large region or country. UNICEF/WHO/UNFPA recommend (as a minimum acceptable level for every 500,000 population) one or more facilities providing comprehensive EOC and four or more facilities providing basic EOC. If areas fall short of this overall minimum level, they may upgrade existing facilities and/or build new ones. If the minimum level is met, evaluators should study the geographical distribution by looking at smaller divisions of the population. National summary measures may hide important sub-national disparities. Disaggregation by geographic (urban/rural) and by administrative (public/private) divisions is recommended (Bertrand and Tsui, 1995).

The use of this indicator in a wide variety of countries has alerted us to at least three difficulties in its application. First, where geographical terrain is particularly challenging and transportation is precarious (such as in the mountains of Nepal or Bhutan), the ratio of facilities to population may require adjustment for local use. Second, the reference period for assessing whether a signal function or procedure has been performed is generally three months, but when patient volume is low, one or more of the signal functions may not be performed, because an occasion did not present itself, not for lack of infrastructure or provider skills. Finally, a third situation concerns normative medical practice that fails to include one of the procedures, for example, assisted vaginal delivery. In some countries, vacuum extraction or a forceps delivery is no longer taught to medical students or midwives and only a few older providers are experienced at performing these procedures.

To solve these problems, one may consider preparing the indicator in several ways. But, to compare facilities across space and time effectively, we recommend maintaining the original operational definitions of these ratios. Evaluators should well document alternative calculations, and should report the adjusted ratio of population to facility; the length of the new reference period, (if it is extended); the way a category of "potential" basic EOC was created (if a procedure is generally performed, but during the study period was not); or the way country-specific criteria were established (if the criteria omits a particular signal function).

Evaluators can also calculate the "number of EOC facilities" for smaller geographical areas to show the distribution of EOC facilities at a sub-national level.

PERCENT OF FACILITIES THAT CONDUCT CASE REVIEW/ AUDITS INTO MATERNAL DEATH/NEAR MISS

Definition

The number of facilities that conduct case review/audits into maternal death/near miss

This indicator is calculated as:

$$\frac{\text{\# of facilities conducting case review/audits into maternal death/near miss}}{\text{\# of facilities at the appropriate level*}} \times 100$$

* Certain facilities will be too small to conduct their own audits, but may participate in established procedures.

Case review refers to a detailed review of the management of a particular patient or "clinical case."

An audit is the systematic and critical analysis of the quality of care. Audit differs from case review because it looks at the whole process of care and at conformity with specified standards of care as part of an iterative cycle of quality improvement (Graham et al., 2000). The different types are as follows:

- Maternal death audits are detailed reviews of the events leading up to a maternal death. The audit may encompass record reviews and staff reports or interviews as well as interviews of relatives/community members;
- Criterion-based audits assess the quality of the clinical management of obstetric complications against defined standards of best practice;
- "Near-miss" audits are performed after the occurrence of a life-threatening event in which a woman is deemed to have nearly died. Criteria for the definition of "near-miss" need clear definitions. (Koblinsky et al., 2000).

Data Requirements

Number of facilities conducting or participating in audits of maternal death and near-miss cases; number of facilities in a specific geographic area

Data Source(s)

Health-facility surveys; district health-management team records

Purpose and Issues

Audit is one of many established mechanisms for improving provider performance in developed countries, and recent studies have shown that it also applies to developing countries.

Evidence of the effectiveness of clinical audit varies, partly because of the differing nature of the interventions assessed (NHS Centre for Reviews and Dissemination, 1999).

This indicator may be collected as part of a facility survey (MEASURE DHS+, 2001), although most programs will need to set up their own monitoring system for assessing the coverage and quality of effective audit practices.

This indicator measures only the proportion of facilities conducting audit or case reviews and does not measure the quality or the impact of the review process. Although case reviews are a routine part of many facility activities, effective audit is not. For this reason, programs may want to collect complementary information on the quality and effectiveness of the process.

In most cases, smaller facilities will find it impractical to conduct their own audit or case reviews. Staff representatives from these facilities, however, should participate in the audit cycles of larger facilities or districts.

PERCENT OF PREGNANT WOMEN ATTENDING ANTENATAL CLINICS SCREENED FOR SYPHILIS

Definition

The percent of pregnant women attending antenatal care screened for syphilis

This indicator is calculated as:

$$\frac{\text{\# of pregnant women attending antenatal clinics screened for syphilis}}{\text{\# of pregnant women attending antenatal clinics}} \times 100$$

This indicator is usually calculated for women attending for their first antenatal visit but may also be collected after delivery.

The most common screening tests for syphilis include rapid plasma reagin (RPR) and venereal disease reference laboratory (VDRL) blood tests.

Data Requirements

The number of women attending antenatal clinics during a reference period (e.g., one year) who were screened for syphilis; the number of women attending the same antenatal clinics during the same reference period

Data Source(s)

Clinic registries (data on first visit) or individual prenatal records (individual ANC records/cards after births or immediately postpartum)

Health facility exit interviews and provider observations are useful for evaluation purposes but not for ongoing monitoring.

Purpose and Issues

The purpose of this indicator is to measure the extent to which ANC clients are screened for syphilis. Since all women attending for ANC should be screened for syphilis at least once during pregnancy, the measure can also potentially serve as a proxy measure of the quality of antenatal care services (UNFPA, 1998a). Furthermore, when an explicit standard exists that all women should be tested at least once during pregnancy, the indicator may also be used as a benchmark to audit provider (or system) performance against compliance with local screening policy.

Syphilis infection is a major cause of maternal morbidity and perinatal morbidity and mortality in the developing world. For many African countries, reported prevalence of syphilis among pregnant women at sentinel surveillance sites ranges between 10-15 percent, with over half these pregnancies resulting in an adverse outcome, such as abortion, stillbirth, low birth weight, premature delivery, or congenital infection (WHO, 1991b). Because adverse outcomes from syphilis are preventable, and screening and treatment in pregnancy are highly cost effective, many countries have adopted universal syphilis screening for pregnant women as a national policy (Gloyd, Chai, and Mercer, 2001).

Screening programs by themselves cannot help reduce the adverse outcomes associated with syphilis and must be linked to efforts to increase ANC coverage and to improve follow up and treatment of women and their partners who test positive.

Researchers may routinely collect data to calculate this indicator if antenatal clinic registries record completed syphilis screening. Most often, however, the information is collected in the context of special surveys that review the antenatal clinic cards of women who have had a recent birth. Researchers may conduct these surveys in facilities or in the community, if women keep their antenatal cards.

Health facility exit interviews and provider observations (MEASURE DHS+, 2001; WHO, 1998a) may provide a baseline measure for evaluation purposes, but are limited because they assess women who have not yet com-

pleted antenatal care and who theoretically could still be tested (MEASURE DHS+, 2001; WHO, 1998a).

The percentage of women screened for syphilis should respond quickly to changes in provider practice, particularly if the indicator is used in a local audit of facility quality of care.

This indicator is a facility-based measure and does not represent the general population, particularly when ANC coverage is low. In addition, where the indicator is obtained by record review, the validity of the findings de-

pends on the quality and completeness of the data. Incomplete data recording may also further indicate low service quality.

Adequate syphilis screening does not equate with adequate syphilis treatment, because studies show that despite effective screening, inadequate treatment can be an important cause of preventable perinatal death. In high prevalence areas, even when syphilis testing is theoretically universal, most women are not tested (Gloyd, Chai, and Mercer, 2001).

PERCENT OF WOMEN WITH OBSTETRICAL COMPLICATIONS TREATED WITHIN TWO HOURS AT A HEALTH FACILITY

Definition:
The percent of women with obstetric complications who are treated within two hours of admittance to a health facility measured during a given reference period

Obstetric complications include:
- Hemorrhage: antepartum, intrapartum or postpartum;
- Prolonged/obstructed labor;
- Postpartum sepsis;
- Complications of abortion;
- Pre-eclampsia/eclampsia;
- Ectopic pregnancy; and
- Ruptured uterus.

Treatment for obstetric complications depends on the complication and local protocols for treatment.

This indicator is calculated as:

$$\frac{\text{\# of women with obstetric complications treated within 2 hours of admittance to a health facility}}{\text{\# of women admitted at a health facility with obstetric complications}} \times 100$$

Data Requirements
Numerator: date and time of admission; date and time of treatment/delivery; total number of women admitted with complications and their diagnosis at the time of admission; and information on the time and nature of treatment given

Denominator: number of women admitted to the health facility with obstetric complications

Data Source(s)
All registers that record where women are admitted to a facility (e.g., labor ward register, antenatal, emergency room or postnatal ward register); registers that record where definitive treatment is administered (e.g., operating theatre register, specific complications registers); case records

Purpose and Issues
The purpose of this indicator is to provide a measure of the quality of maternity care, because maternal mortality is directly related to the effectiveness and timelines of treatment for emergency complications (Koblinsky, et al., 1995).

This indicator is most appropriate at a facility level for those facilities interested in auditing their own care practice. It is less suitable for comparison purposes across facilities because of variation in the services provided and in case mix at different facilities as well as variations in the definitions of the indicator.

Information required to construct the indicator should be available directly from facility registers and case notes. The lack of certain information may signal suboptimal care/management. The feasibility of the indicator also critically depends on standard definitions of admission-to-treatment-time-interval (ATTI) for each obstetric complication. For example, does admission time mean time of first arrival at the hospital, time seen by the admissions clerk, or something else? What is the treatment time for an obstetric hemorrhage? Is it when an intravenous drip is inserted, when a blood transfusion starts, when an oxytoxic is given, or when the hemorrhage stops?

This indicator should respond rapidly to changes in staff or facility administration practice. Early rapid response may simply be due to a Hawthorne effect (i.e., apparent improvement simply because of an observation taking place), but this type of improvement will not be sustained. To maintain improvement, programs should regularly audit ATTI. Better still, ATTI could be incorporated into routine management practice (e.g., discussed routinely in all case reviews). (See indicator, **Percent of Facilities that Conduct Case Reviews/Audits into Maternal Death/Near Miss**).

This type of indicator has several ambiguities and deficiencies, as follows:

- The recording of the actual admission time may be seriously delayed if the facility lacks a triage system, and women must wait a long time on first arriving at the facility;
- The ATTI should be reviewed for all women developing obstetric complications. If the admissions register serves as the sampling frame to identify women with complications, then the register will exclude those women who had a complication but whose diagnosis was missed at the time of admission;
- The severity of the complication and hence the need for rapid treatment may be difficult to ascertain retrospectively for all cases (a further reason for facilities to set their own realistic standards);
- The indicator fails to capture the timeliness of treatment for those women who develop complications while in the hospital;
- Women who arrive at the hospital, but who are not admitted for a variety of reasons (e.g., they cannot afford the treatment), will be omitted from both the numerator and denominator. This omission may be a potentially important source of bias if poorer patients have a disproportionately high complication rate; and
- The indicator does not measure the appropriateness of the type of treatment given.

Because of these limitations, evaluators can best measure ATTI in combination with other indicators or approaches that measure complementary aspects of the quality of care, for example in the context of care reviews or near-miss audits.

Indicator

CESAREAN SECTIONS AS A PERCENT OF ALL LIVE BIRTHS

Definition

The percent of pregnant women who have a cesarean section in a specific geographical area and reference period

This indicator is calculated as:

$$\frac{\text{\# of cesarean sections performed}}{\text{\# of live births}} \times 100$$

Data Requirements

The number of cesarean sections performed in a defined population during a reference period; total number of live births in the same reference period

Data Source(s)

Numerator: clinical registries for data in a given geographical area on the number of C-sections performed; estimates of the number of births in that area; and population-based surveys for self-reported C-sections only

Denominator: all live births during the reference period. Where data on the numbers of live births are absent, evaluators can calculate total estimated live births using census data for the total population and crude birth rates in a specified area. *Total expected births = population x crude birth rate.*

Household demographic surveys often produce national and disaggregated estimates of the self-reported C-section rate.

Purpose and Issues

This indicator demonstrates the extent to which a particular life-saving obstetric service is being performed in EOC facilities. It reflects the accessibility and utilization of services as well as the functioning of the health service system. The appropriate use of a cesarean section leads to a decrease in maternal mortality and morbidity, as well as a decrease in perinatal morbidity and mortality. While cesarean sections may be performed solely for the health of the fetus or newborn, in developing countries the vast majority relate to maternal indications.

Many of the major pre- and intrapartum causes of maternal mortality and morbidity require the use of this procedure to save the woman's life or to prevent serious morbidity.

Of all the procedures used to treat the major obstetric complications, C-sections may be the easiest to study because record-keeping for C-sections is more reliable than that for other procedures or obstetric complications (MotherCare, 2000b; UNICEF, WHO, UNFPA, 1997). However, it is critical that evaluators include information for all facilities performing C-sections in the area under study in the numerator.

Changes in the ability of the health care system to provide cesarean sections can have an impact within six to nine months.

UNICEF/WHO/UNFPA recommend a C-section rate between 5 and 15 percent of all births, based on estimates from a variety of sources. Rates less than 5 percent may indicate inadequate availability and/or access to EOC. Rates above 15 percent suggest overuse of the procedure for non-emergency reasons. Excessive use unnecessarily exposes women to anesthesia and surgery with their concomitant risks. Moreover, it drains scarce health-care resources. Most of the countries with excessively high C-section rates are also highly litigious societies such as the United States, where 22 percent of all births are cesareans (Lancet, 2000). However, Brazil has a rate of at least 36 percent of all live births (BEMFAM and Macro International, 1997).

Disaggregation of the rate allows one to evaluate access to the procedure. Rates are often inconsistent between urban and rural environments, public or private sectors, different payment schemes, or across regions. Thus, sub-national estimates are encouraged (Maine, McCarthy, and Ward, 1992).

Crude birth rates produce estimates of live births only, whereas some cesarean sections are performed on pregnancies that result in stillbirths. If the number of C-sections performed for stillbirths is low, the use of live births should be acceptable as the denominator.

An alternative indicator, the proportion of facility deliveries that are C-sections, will vary by the case mix of patients and will be biased by referral patterns of women with complications requiring the procedure. Specifying an appropriate range of target percentages within a facility is impractical.

The procedure of cesarean section usually occurs at the end of a complex series of events, possibly including pre-existing and pregnancy-specific medical factors, identification of complications, transportation to health-care facilities and availability of necessary technology. When using this indicator, managers and evaluators may also want to employ more in-depth techniques, such as case audits, to investigate what clinical indicators are being used for cesarean section and if the appropriate women are receiving this service. By itself, the indicator reveals nothing about the appropriateness of the procedure.

Case Fatality Rate (CFR) – All Complications[4]

Definition

The proportion of women with major obstetric complications who die in a facility within a reference period

This indicator is calculated as:

> # of deaths from specified obstetric
> complications in a facility
> ─────────────────────────────── x 100
> # of women with specified obstetric
> complications attended in the facility

Where deaths from the following complications are included:

- Hemorrhage: antepartum, intrapartum or post-partum;
- Prolonged/obstructed labor;
- Postpartum sepsis;
- Complications of abortion;
- Pre-eclampsia/eclampsia;
- Ectopic pregnancy; and
- Ruptured uterus.

All cases in the numerator also appear in the denominator. All complications specified in the list above appear in both the numerator and denominator. By definition, a CFR is cause-specific, but in this case, a single facility may only see a small number of women with any one complication.

Data Requirements

The number of deaths from the specified complications in the facility during the specified time period; the number of women diagnosed with one or more of these complications attended at the EOC facility during the specified time period

Data Sources

Facility records

Purpose and Issues

This indicator measures facility performance, in particular, quality and promptness of care. It is most useful when comparisons are made over time for the same facility. It is not useful for comparisons across facilities of different types because of variations in the characteristics of the client populations and in the services provided by the facilities. Women with more severe complications are more likely to present at referral hospitals, whereas women with less severe complications may access district hospitals or health centers (MotherCare, 2000a). Even comparisons among "same level" or "like" facilities may be difficult to interpret, as the population profile can vary dramatically because of socio-cultural factors or other circumstances outside the control of the health sector, such as transportation and road systems.

The CFR has an extremely strong causal link to maternal mortality at the facility level. Its relationship to maternal mortality in the general population depends on the proportion of women with obstetric complications who are managed in facilities. The higher the number of these women managed in facilities is, the closer the relationship between CFRs and the level of maternal mortality in the general population (Bertrand and Tsui, 1995).

If the facility treats obstetric complications and collects data on obstetric complications and maternal deaths, this indicator is easy to calculate. Case fatality rate should respond to changes within a fairly short period of time (e.g., 6-12 months).

Whether a woman dies in the hospital will depend not only on the quality and readiness of the hospital's response to a woman with an obstetric emergency, but

---

[4] Much of the text for this indicator comes from Maine, McCarthy, and Ward, 1992 and UNICEF, WHO, and UNFPA, 1997.

also on her condition on admission to the hospital. Thus, the hospital could function well and still have a high CFR because women in need of EOC arrive in such poor condition. Where a facility's CFR is low, the quality of care is not necessarily high; instead, few women with obstetric complications may use services. For these reasons, one should have other indicators of quality of care (e.g., the time interval between admission and treatment for women with complications) or more in-depth information on the woman's status at admission (e.g., pulse, blood pressure, and temperature).

Finally, this indicator is helpful with the other four to five "program-level" indicators that UNICEF/WHO/ UNFPA recommend. For example, if the percentage of all births in EOC facilities or met need is low, the CFR may be relatively meaningless.

UNICEF/WHO/UNFPA recommend a maximum acceptable value for the indicator of less than 1 percent. However, a study of US hospitals showed a CFR of 0.03 percent in 1978 (Petitti et al., 1982). Certainly, countries meeting even a level of 1 percent should strive to reduce the rate.

Where the number of maternal deaths or complicated cases is small, the CFR will not be sufficiently robust to be meaningful. However, where the number of cases is large, evaluators can calculate CFRs for individual complications.

PERCENT OF AUDIENCE THAT KNOW THREE PRIMARY WARNING/DANGER SIGNS OF OBSTETRIC COMPLICATIONS

Definition
Community knowledge and awareness of the warning/danger signs of obstetric complications

"Audience" is the intended population for the program (e.g., pregnant women, husbands or other members of the community who influence decisions about care seeking at the time of delivery).

"Know" refers to the percentage who can spontaneously name at least three primary warning signs of specific obstetric complications, which in a wide range of settings include:

- Antepartum bleeding;
- Labor >12 hours (obstructed labor);
- Placenta retained > 1 hour;
- Convulsions or fit or swelling of the hands or face (pre-eclampsia/eclampsia); and
- Fever and vaginal discharge (puerperal sepsis)[Filippi et al., 2000].

This indicator is calculated as:

$$\frac{\text{\# of audience who know at least three warning/danger signs of obstetric complications}}{\text{Total \# in the intended audience}} \times 100$$

Data Requirements
Number of the audience who can name at least three of the warning signs of obstetric complications (listed above); total population defined as the intended audience

Data Source(s)
Population-based surveys

Purpose and Issues
The purpose of this indicator is to assess community knowledge and awareness of the warning/danger signs of obstetric complications in order to plan and monitor the impact of BCC initiatives at a community level.

Knowledge of the danger signs of obstetric complications is the essential first step in the appropriate and timely referral to essential obstetric and newborn care services (Perreira et al., 2001).

Improvement in knowledge of obstetric complications is usually much smaller than improvements in other health-education messages such as self care (MotherCare, 2000a and 2000b).

Knowledge of the danger signs of an obstetric complication is only one aspect of obstetric-problem recognition at the community level. Knowledge about the severity of an obstetric complication (i.e., knowing when to take action) and knowledge about the appropriate life-saving action for each complication are also important. Moreover, adequate knowledge does not guarantee that an individual will recognize a complication in practice. Certain obstetric complications that evolve from a normal to an emergency state (e.g., postpartum hemorrhage) may be particularly difficult to recognize. Care seeking is also strongly influenced by cultural beliefs about the etiology of an illness. These beliefs may more powerfully influence an individual's action than will his/her recent knowledge of the appropriate action to take (MotherCare, 2000a and 2000b).

Evaluators should combine knowledge of pregnancy danger signs with other indicators measuring related aspects of knowledge and behavior to assess the real impact of any BCC program. Complementary indicators can include, for example, the percentage of the population who know the location of emergency obstetric services and the percentage of the population who intend to use these services in the event of emergency (See Part II.F).

Indicators of knowledge of danger signs and related indicators should be complemented by good quality formative research and, where appropriate, qualitative methodologies such as illness narratives (MotherCare, 2000a and 2000b).

Gender Implications of this Indicator

Most maternal deaths are due to sudden and unexpected complications. To reduce the nearly 600,000 deaths occurring each year from largely preventable causes, much effort has focused on training health workers and pregnant women to recognize danger signs so that serious complications are recognized soon enough to receive medical attention. Few efforts have been made to educate men about the risks of pregnancy, even though men often control decisions to seek medical attention and often arrange and pay for transport to a health facility. If men as well as women understood that all pregnancies carry some risk, complications would be recognized and treated.

Percent of Women Attended at Least Once During Pregnancy for Reasons Related to the Pregnancy

Definition

The percent of women attended at least once during pregnancy by skilled health personnel for reasons related to the pregnancy

This indicator is calculated as:

$$\frac{\text{\# of pregnant women attended by skilled personnel at least once during their pregnancy for reasons related to the pregnancy}}{\text{Total \# of live births occurring within the reference period}} \times 100$$

A skilled attendant is a professional care giver who possesses the knowledge and a defined set of cognitive and practical skills enabling the individual to provide safe and effective health care during childbirth to women and their infants in the home, health center, and hospital settings. Skilled attendants include midwives, doctors, and nurses with midwifery and life-saving skills[5] (Inter-Agency Group for Safe Motherhood, 2000).

Data Requirements

Numbers of women who are seen by skilled personnel during pregnancy; all live births in a reference period

The number of live births is a proxy for the numbers of all women who need antenatal care (ANC). Evaluators should include all births, but they usually use only live births because of the difficulty in obtaining information about non-live births (Graham and Filippi, 1994).

Where data on the numbers of live births are unavailable, evaluators can calculate total estimated live births using census data for the total population and crude birth rates in a specified area. *Total expected births = population x crude birth rate*

Data Source(s)

Routine health services data, population-based surveys

Routine health service data typically lack information on pregnancies or births that take place outside the public health sector, for example in homes or private facilities.

Purpose and Issues

The main purpose of an indicator of antenatal care is to provide information on women's use of antenatal care services. ANC coverage provides a crude measure of antenatal care utilization (Rooney, 1992), but it does not capture the number and timing of visits, the reasons for seeking care, the skills of the provider, or the quality of care received. Therefore, evaluators should not infer that similar rates of ANC coverage mean similar levels of care.

Although epidemiological studies tend to show an association between improved maternal health outcome and ANC, most fail to control for selection biases that would positively influence the outcome (Villar and Khan-Neelofur, 2000). The association between one antenatal visit (with care provision of unknown quality) and maternal mortality is weak (WHO, 1999b). However, women's use of ANC is more strongly associated with improved perinatal survival (McDonagh, 1996); therefore, measures of ANC coverage may have a greater role in the monitoring and evaluation of programs addressing newborn health and survival (Graham and Filippi, 1994).

This indicator is responsive to change in the short term. Annual monitoring is only feasible when the data are derived from routine data sources. For international comparisons, periods of three to five years are probably sufficient. Evaluators should avoid frequent surveys, because sampling error makes it difficult to assess whether small changes are real or due to chance variation.

---

[5] A trained traditional birth attendant is NOT included in the definition of skilled personnel

For comparison purposes, one must know whether the denominator used reflects all births, the most recent birth, or all women. A birth-based analysis represents all births in the survey period but over-represents women who have more than one birth. Women with more than one birth are also more likely to have other risk factors, such as high parity and lower rates of health services use. Hence, ANC coverage is likely be lower using a birth-based estimate than a woman-based estimate, and this difference will be greater the longer the survey period used. One can obtain a women-based estimate by using ANC coverage for the most recent birth (Graham and Filippi, 1994); this format is now used in DHS reports. Because programs target women, using a woman-based denominator may be conceptually more appealing to program managers. However, a birth-based analysis for all births (live births and stillbirths) is essential for determining the impact of ANC on pregnancy outcomes.

Differences in the categorization of skilled personnel, in particular whether auxiliary staff or traditional birth attendants (TBA) have been included, may also account for discrepancies between countries. In practice, most large surveys now use the standard WHO definition of skilled health personnel.

Discrepancies may arise because the estimate relates either to all antenatal visits or only to those occurring "for reason related to the pregnancy." In practice, information on women's care-seeking motives is rarely collected except in Pan Arab Family Health Surveys (PAPFAM).

Antenatal care coverage is one of four mutually supportive indicators in the minimal list measuring maternal health service coverage. The other three indicators are:

- **Percent of Births Attended by Skilled Health Personnel;**
- Availability of basic essential obstetric care; and
- Availability of comprehensive essential obstetric care.

In combination, these indicators measure progress towards the goal of providing antenatal care, trained attendants during childbirth, and access to essential obstetric care for all pregnant women. ANC coverage is associated with newborn health and survival and weakly associated with maternal mortality. In sum, antenatal care coverage appears to influence newborn health and survival, but its effect on maternal mortality is unclear.

Gender Implications of this Indicator

Because some countries deem it culturally inappropriate for women to discuss issues related to their bodies with men, women may not be able to communicate pregnancy-related problems to male providers. In addition, where women lack access to household resources or where they lack the autonomy to seek health care on their own, husbands or other family members may not be willing to invest resources in antenatal care, particularly if a given pregnancy is progressing "normally."

PERCENT OF WOMEN WHO WERE GIVEN OR PURCHASED MALARIA PROPHYLAXIS/TREATMENT DURING THEIR MOST RECENT PREGNANCY

Definition

The percent of women who were given or who purchased malaria medication, according to national policy, during their most recent pregnancy

This indicator is calculated as:

$$
\frac{\text{\# of women who were given or purchased malaria medication during their most recent pregnancy}}{\text{\# of women who had a recent live birth}} \times 100
$$

Malaria medication (prophylaxis or presumptive intermittent treatment [PIT]) will vary according to local susceptibility and national policy. Most country policies in highly endemic areas advise a treatment dose of chloroquine followed by weekly chloroquine throughout pregnancy or PIT with sulphadoxine-pyrimethamine at the beginning of the second and third trimester in areas of chloroquine resistance. Other regimes are less common.

Data Requirements

Number of women who were given or who purchased malaria medication during their most recent pregnancy; number of women with a recent live birth

For both the numerator and denominator, evaluators should specify the time periods for the recent live birth. In surveys, this period is normally restricted to three to five years before the survey.

Where data on the numbers of live births are absent, evaluators can calculate total estimated live births using census data for the total population and crude birth rates in a specified area. *Total expected births = population x crude birth rate*

Data Source(s)

Population-based surveys; facility records of antenatal patients

Purpose and Issues

Malaria poses a serious health risk to pregnant women and newborns, particularly in areas where *p. falciparum* is endemic. Malaria is a major cause of maternal anemia and of low birth weight, (independently, low birth weight is the single most important determinant of neonatal mortality). It is often the primary cause of intrauterine growth retardation in the absence of severe maternal malnutrition. As research demonstrates, both PIT with sulphadoxine-pyrimethamine and routine chemoprophylaxis with chloroquine reduce these complications, although the evidence is weaker for chloroquine because of worsening drug resistance. In 2000, the African summit, "Roll Back Malaria," established a goal that at least 60 percent of women at risk of malaria should have access to chemoprophylaxis or PIT.

Some large household surveys, such as in the DHS core questionnaire, routinely collect data for this indicator. In addition, some health facility surveys that conduct record reviews, direct observation of ANC consultations, or exit interviews with ANC clients yield this information for client populations. Because malaria varies within communities, evaluators should disaggregate population-based data by area, where possible. This will also help to monitor drug resistance.

Many facilities routinely record tetanus-toxoid coverage and numbers of ANC visits, and HIS could collect this information with little additional effort. However, as far as we are aware, routine collection of this indicator only occurs in the context of special studies.

One major limitation with the indicator is that current data collection approaches lack information on the completeness of the drug regimen taken during pregnancy. In addition to determining the type of malaria medication taken, information on the frequency and timing of drug administration is required to determine whether pregnant women are adequately protected against malaria.

Information on the frequency and timing of drug administration could theoretically be obtained if clinics maintained records on the numbers of patients attending and on the number of women given first, second, and third courses of PIT or the number of packets of chloroquine dispensed. An alternative indicator reflecting the adequacy of the program in meeting the needs of specific clients is

- Number of tablets distributed per eligible woman.

The questions asked in most population-based surveys assume that women are able to report on malaria treatment reliably, but few validation studies have tested this assumption. Population-based studies also rely on self-reported data, which are subject to recall bias likely to increase with the length of the recall period.

Facility records or provider observation measures the proportion of women given or prescribed malaria medication but does not reflect the proportion of women who took the medication. Compliance with treatment will rarely be 100 percent and will vary depending on many different local factors.

Where malaria is sporadic or seasonal, programs focus on screening women who present with symptoms and on treating those who are infected. Alternative indicators in this case include

- Number of pregnant women presenting with malaria symptoms; and
- Percent of pregnant women treated for malaria according to locally established protocols.

Alternative indicators for both endemic and sporadic areas are:

- Percent of pregnant women living in a household with treated mosquito nets; and
- Percent of pregnant women living in a household with treated mosquito nets that report having slept under the net the previous night.

PERCENT OF PREGNANT WOMEN WHO RECEIVE ANTIHELMINTHIC TREATMENT DURING PREGNANCY

Definition

In areas of moderate to high endemicity: the percent of pregnant women who receive presumptive antihelminthic treatment during their pregnancy. According to the 1998 IVACG/WHO/UNICEF "Guidelines for the Use of Iron Supplements to Prevent and Treat Iron Deficiency Anemia," treatment should be done once in the second and third trimester.

In areas of low endemicity: the percent of pregnant women who received prescribed treatment during their pregnancy.

This indicator is calculated as:

$$\frac{\text{\# of pregnant women who receive presumptive antihelminthic treatment}}{\text{Total \# of pregnant women}} \times 100$$

Data Requirements

Information on the number of pregnant women who receive presumptive/prescribed anthelminthic treatment and the total number of pregnant women.

Data Source(s)

Program records (on number of pregnant women, number of pregnant women who receive treatment, either presumptive treatment or therapy for identified helm-inths, and number of pregnant women reported to be infected); population-based surveys represent an alternative source of data, but will yield different results in terms of coverage. If the source of data is a population-based survey, the evaluator should calculate the indicator for the last pregnancy.

Purpose and Issues

Helminths such as hookworm and schistosomes can cause blood and iron loss and thus anemia. In areas of low endemicity, treatment in the second trimester is recommended. In areas of moderate to high endemicity, treatment should occur in the second and third trimester. Treatment in the first trimester is not recommended.

This indicator only measures whether women have received any anthelminthic therapy, without reference to adequate dosing. Because treatment of helminths depends on the availability of medication to clients in the program, this indicator may reflect inadequacies in the flow of drugs to service distribution points in the system and/or poor provider performance at the SDP.

An alternative indicator reflecting the adequacy of the program in meeting the needs of specific clients is the dosage (number of tablets) distributed per eligible woman.

PERCENT OF BIRTHS ATTENDED BY SKILLED HEALTH PERSONNEL

Definition
The percent of births attended by skilled health personnel

This indicator is calculated as:

$$\frac{\text{\# of births attended by skilled personnel during the reference period}}{\text{Total \# of live births occurring within the reference period}} \times 100$$

The skilled attendant is a professional care giver who possesses the knowledge and a defined set of cognitive and practical skills that enable the individual to provide safe and effective health care during childbirth to women and their infants in the home, health center, and hospital settings. Skilled attendants include midwives, doctors, and nurses with midwifery and life-saving skills[6] (Safe Motherhood Inter-Agency Group, 2000).

Data Requirements
Number of births attended by skilled health personnel in a defined time period; number of live births in the same geographic area and reference period

The number of live births is a proxy for the numbers of women who need delivery care. Evaluators should count all births but usually only use live births in calculating this indicator, because of the difficulty in obtaining information about non-live births (Graham and Filippi, 1994).

Where data on the numbers of live births are unavailable, evaluators can calculate total estimated live births using census data for the total population and crude birth rates in a specified area. *Total expected births = population x crude birth rate*

Data Source(s)
Routine health service data; population-based surveys

Routine health service data typically lack information on pregnancies or births that take place outside the public health sector, for example in homes or private facilities.

Purpose and Issues
The main purpose of an indicator of the skilled attendant at delivery is to provide information on women's use of delivery care services.

Many argue that increasing the proportion of deliveries with a skilled attendant is the single most critical intervention for reducing maternal mortality. Moreover, the proportion of births with a skilled attendant is a benchmark indicator for monitoring progress towards the ICPD goals (WHO, 1999b).

The evidence that delivery with a skilled attendant reduces maternal mortality comes from a number of clinical, historical, and epidemiological sources that indicate an association but not a causal link. In general, births with a skilled attendant are associated with lower rates of maternal mortality. However, confounding factors, such as the strong correlation between skilled attendant and institutional delivery, make assessing the impact of skilled attendant alone difficult to determine.

Annual monitoring is only feasible when the data are derived from routine data sources. For international comparisons, periods of three to five years are probably sufficient. Frequent surveys are generally undesirable because the survey periods may overlap, and sampling error makes it difficult to assess whether small changes are real or due to chance variation.

Evaluators should not infer that similar rates of skilled attendant deliveries between countries reflect similar levels of care; major differences are likely to exist between countries in how providers are trained, in what providers are allowed to practice and do practice, and

---

[6]A trained traditional birth attendant is NOT included in the definition of skilled personnel.

in what resources, equipment, and supplies are at their disposal.

Differences in what definitions are used and in how skilled attendants are reported may also account for discrepancies between countries. Most surveys such as the DHS rely on women's self-report but how women interpret the question "who assisted with the delivery?" and whether they accurately identify the health staff attending is unknown.

This indicator uses a birth-based analysis (similar to the previous ANC indicator), and the sample will over-represent women with multiple births in the survey period. Women with more than one birth are also more likely to have other risk factors, such as high parity and lower rates of health services use. Delivery coverage may therefore be underestimated, although this underestimate is likely to be small.

Since the denominator for this calculation includes only women with live births and excludes women with fetal deaths and stillbirths, the only valid association will be with neonatal mortality and not with perinatal mortality. (See Part III.E Newborn Health.)

Evaluators can disaggregate skilled attendant at delivery by place to further document the degree of care received at the time of delivery. This measure of care or "skilled attendance" will vary by setting and attendant. A skilled attendant conducting a delivery in hospital, for example, provides a higher level of "skilled attendance" than does a skilled attendant conducting a delivery at home.

The percentage of births with a skilled attendant is one of four mutually supportive indicators in the minimal list measuring maternal health services coverage. The other three indicators are:

- **Percent of Women Attended at least Once during Pregnancy for Reasons Related to the Pregnancy;**
- **Availability of basic essential obstetric care; and**
- **Availability of comprehensive essential obstetric care.**

In combination, these indicators measure progress towards the goals of providing antenatal care, trained attendants during childbirth, and access to essential obstetric care for all pregnant women.

Percent of Women Attended During the Postpartum Period by Skilled Personnel

Definition

The percent of pregnant women seen during the postpartum period by a skilled attendant

This indicator is calculated as:

$$\frac{\text{\# of women attended during the postpartum period by skilled personnel}}{\text{Total \# of live births}} \times 100$$

The postpartum period is defined as the time from the delivery of the placenta until 42 days after delivery.

Data Requirements

Number of women within the postpartum period who are attended by skilled personnel during a fixed time period (Note: One must specify whether the visit was a first or subsequent visit); all live births during the same time period

The number of live births is a proxy for the numbers of all women who need postnatal care. Evaluators generally count all births, but usually use only live births to calculate this indicator because of the difficulty in obtaining information about non-live births (Graham and Filippi, 1994).

Where data on the numbers of live births are unavailable, evaluators can calculate total estimated live births using census data for the total population and crude birth rates in a specified area. *Total expected births = population x crude birth rate*

Data Source(s)

Routine health service data; population-based surveys

Routine health service data typically lack information on pregnancies or births that take place outside the public health sector, for example in homes or in private sector facilities.

Purpose and Issues

The main purpose of an indicator for postpartum care is to provide information on women's use of postpartum services in the postpartum period.

Although most maternal and newborn deaths occur around the time of delivery and in the immediate postpartum period, postpartum care (PPC) has been a relatively neglected area of maternity-services provision. Recent WHO guidelines recommend that the first postpartum visit take place within the first week, preferably within the first two to three days, with a second visit at four to six weeks. The visit should include the early detection and treatment of complications and preventive care for both mother and baby[7] (WHO, 1998a; WHO, 2001b).

Because pregnancy complications occur unpredictably, the likelihood is low that one postpartum visit of unspecified content, quality, and timing may influence maternal mortality. A postnatal care visit may reduce the likelihood of severe morbidity if pregnancy complications are detected early and if effective treatment is available (Child Health Research Project, 1999). Further research in this area is required.

Some large surveys such as DHS routinely collect data on postpartum care. Routine HIS may also collect data, although historically more programs have collected data on postnatal care for the baby (for immunization coverage) than for the mother.

This indicator is responsive to change in the short term. Annual monitoring is only feasible when the data are derived from routine data sources. For international comparisons, periods of three to five years are probably sufficient. Frequent surveys are probably undesir-

---

[7] Preventive care for mothers may include vaccination against tetanus; provision of vitamin A and iron; and counseling on appropriate newborn care, hygiene, breastfeeding, malaria prevention and nutrition. Preventive care for babies may include early BCG, and polio and hepatitis immunization.

able because sampling error makes it difficult to assess whether small changes are real or due to chance variation.

Postpartum care is a package of services instead of one single intervention. Because no widely accepted operational definition of postpartum care exists and because the content and quality of care are likely to vary between settings, similar coverage rates between countries do not reflect similar levels of care.

Furthermore, after delivery, the two individuals need very different care and attention. Postpartum care statistics should make explicit whether care was provided principally for the mother or baby or both mother and baby, because this information may be difficult to determine retrospectively. The current DHS questionnaire, for example, specifies postpartum care for the mother. WHO distinguishes between care for the mother and for the baby by using the term postpartum to refer to care exclusively for the mother and the term postnatal care for the baby.

Some surveys present only postpartum coverage for women who delivered outside a facility on the assumption that women who deliver in facilities receive some degree of postpartum care (Rutstein, 1999).

In settings where postnatal coverage is relatively high, evaluators can stratify this indicator by time of visit (e.g., within two or three days, one week or later) to better measure women's use of services.

MATERNAL MORTALITY RATIO (MMR)

Definition[8]
The number of maternal deaths per 100,000 live births

A maternal death (as cited in International Classification of Disease or ICD-10, [WHO, 1992]) is the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and the site of the pregnancy. Death can stem from any cause related to or aggravated by the pregnancy or its management, but not from accidental or incidental causes. Maternal deaths fall into two groups, direct and indirect, as follows:

Direct obstetric deaths
Direct obstetric deaths result from obstetric complications of the pregnant state (pregnancy, labor, and puerperium), from interventions, omissions, incorrect treatment, or from a chain of events resulting from any of the above.

Indirect obstetric deaths
Indirect obstetric deaths result from previous existing disease or disease that developed during pregnancy and which was not due to direct obstetric causes, but which was aggravated by physiologic effects of pregnancy.

The maternal mortality ratio is calculated as:

$$\frac{\text{All maternal deaths occurring within a reference period (usually 1 year)}}{\text{Total \# of live births occurring within the reference period}} \times 100,000$$

Data Requirements
Information on all maternal deaths occurring in a period (usually 1 year) and information on the total number of live births occurring in the same year

Including all pregnancies in the denominator gives a true indication of the total population of pregnant and delivering women at risk of maternal death, but

researchers and evaluators more commonly use live births since these data are more readily available and are easier to collect.

Where data on the numbers of live births are absent, evaluators can calculate total estimated live births using census data for the total population and crude birth rates in a specified area. *Total expected births = population x crude birth rate*

Data Sources:
Many countries have three main sources of data with which to calculate the maternal mortality ratio:

- Vital registration;
- Service statistics; and
- Population based surveys or surveillance.

The serious limitations of these sources, in both the developing and developed world, have been well documented elsewhere. (AbouZahr, 1999; Berg, Danel, and Mora, 1996; Campbell and Graham, 1990).

Purpose and Issues
The maternal mortality ratio is the most widely used measure of maternal deaths. It measures obstetric risk (i.e., the risk of dying once a woman is pregnant). It therefore omits the risk of being pregnant (i.e., fertility, in a population, which is measured by the maternal mortality rate or the lifetime risk) (Graham and Airey, 1987).

Maternal mortality is widely acknowledged as a general indicator of the overall health of a population, of the status of women in society, and of the functioning of the health system. High maternal mortality ratios are thus markers of wider problems of health status, gender inequalities, and health services in a country. The maternal mortality ratio is therefore useful for advocacy

---

[8] The text for this indicator draws heavily on the forthcoming WHO publication, "Indicators for Reproductive Health Monitoring," WHO.

purposes, but lacks information on the causes of high maternal mortality or the interventions required to reduce maternal deaths.

Maternal deaths are difficult to investigate because of their comparative rarity on a population basis, as well other context-specific factors, such as reluctance to report abortion-related deaths, problems of memory recall, or lack of medical attribution (Campbell and Graham, 1991). Thus, no single source or data collection method is adequate for investigating all aspects of maternal mortality in all settings.

Few developing countries have vital registration systems sufficiently complete to provide reliable population estimates (AbouZahr, 1998).

The main drawback of health services data relates to the selectivity of the service-using population. Without detailed knowledge of the catchment population, it is hard to gauge whether the maternal mortality ratio under or over estimates the level for the general population (which also includes non-users of the service). Other problems related to using health services information include inaccuracies in routine registers and omission of deaths occurring outside maternity wards.

Population based surveys are the primary source of information for calculating the maternal mortality ratio in many developing countries. These types of surveys include:

- **RAMOS** (Reproductive Age Mortality Surveys) studies seek to identify all female deaths in the reproductive period, using a combination of approaches, such as cross-sectional household surveys, continuous population surveillance, hospital and health center records, and key informants (WHO, 1987).

- **Direct estimation** relies on asking questions about maternal deaths in a household during a recent interval of time, say one to two years. These questions can be asked in the context of a household survey or a census of all households, although as yet experience with the latter is fairly limited (Campbell, 1999).

Both these types of methods provide up-to-date estimates but are time-consuming and costly because they require large sample sizes to obtain single-point estimates with sufficiently narrow confidence intervals to enable monitoring of time trends.

- **The sisterhood method** goes some way to overcoming large sample size requirements by interviewing adult respondents about the survival of all their sisters. The indirect method (Graham, Brass, and Snow, 1989) involves fewer questions to respondents but provides a pooled estimate that relates statistically to a point around 10-12 years prior to the survey. The direct method (Stanton, Abderrahim, and Hill, 2000) provides a more current estimate at about 3-4 years prior to the survey, but requires more questions and is more costly and time consuming.

Because of the imprecision in these estimates, modeling methods have also been developed, (WHO, UNICEF, and UNFPA, 2001; AbouZahr and Wardlaw, 2001; UNFPA, 1998b).

Maternal mortality ratios are only a broad indication of the level of maternal mortality, rather than a precise measure, because of the limitations inherent in most measurement methods. The use of confidence intervals around the estimates helps raise awareness that a point estimate is usually too imprecise to be used to monitor trends (AbouZahr and Wardlaw, 2001). Furthermore, the data sources and collection methods described above have very different strengths and weaknesses and yield estimates of varying reliability. For this reason, surveys to estimate maternal mortality should occur no more frequently than every 5-10 years. Evaluators must take into account the large confidence intervals in interpreting the maternal mortality ratio.

Distinguishing between real and artificial changes in the maternal mortality ratio is complicated because observed differences do not necessarily indicate improved maternal health status (Graham, Filippi, and Ronsmans, 1996). Other important issues to consider include:

- Non-sampling errors such as changes in the accuracy of reporting or of classification over time or between districts or populations (Stanton, Abderrahim, and Hill 2000);
- Changes in the definition of a maternal death between ICD-9 and -10 (WHO, 1977; WHO,

1992). Presentation of the maternal mortality ratio should thus clearly state which version it used. In the case of ICD-10, one must specify which of the three categories (direct and indirect maternal deaths up to 42 days postpartum, late maternal deaths, pregnancy-related deaths[9]) the numerator includes;

- Aggregate levels may hide wide differentials between population subgroups; and
- Apparent differences in the maternal mortality ratio between rural and urban areas may simply reflect differences in the pattern (not level) of fertility, with more rural women who are grand multiparous and for whom the risk of death will likely be higher. Other possible confounders include general health status, such as levels of anemia or malaria, and socio-economic factors.

---

[9] Late maternal deaths: direct or indirect obstetric causes more than 42 days but less than one year after the termination of pregnancy. Pregnancy-related deaths: deaths while pregnant or within 42 days of the termination of pregnancy, irrespective of the cause.

Met Need for Essential Obstetric Care (EOC)[10]

Definition

The percent of all women with major obstetric complications who are treated in EOC facilities in a given reference period

This indicator is calculated as:

> # of women with major obstetric complication treated in EOC facilities
> ───────────────────────────────── x 100
> Estimated # of women with obstetric complications from the geographical area served by the EOC facilities

Where the direct or major obstetric complications include:

- Hemorrhage: antepartum, intrapartum, or postpartum;
- Prolonged/obstructed labor;
- Postpartum sepsis;
- Complications of abortion;
- Pre-eclampsia/eclampsia;
- Ectopic pregnancy; and
- Ruptured uterus.

Number of women with a major obstetric complication includes both women admitted with the complication and women who develop the complication in the facility.

EOC facilities include both "basic" and "comprehensive" levels of essential obstetric care.

Data Requirements

The number of women with a major obstetric complication treated in EOC facilities during the reference period; (an estimate of) the number of women with major obstetric complications in the population during the reference period

Data Sources

Facility records (for number of women treated)

The number of pregnant women who develop obstetric complications requiring medical care to avoid death or disability is estimated to be 15 percent (WHO, 1994a). The number of live births frequently serves as a proxy for all births or pregnancies; when data on the numbers of live births are absent, evaluators can estimate them from *total expected births = population x crude birth rate*.

Purpose and Issues

The purpose of this indicator is to gauge the level of use of EOC services by women experiencing a major obstetric complication in a specified time period and geographical area.

Facility record-keeping systems may require adjustments for the routine collection of data on obstetric complications. A useful system will record major complications in the patient register or maternity logbook. Evaluators must ensure that they gather information from all relevant parts of the facility (e.g., gynecology ward, surgical ward, abortion ward, morgue) and not just from the maternity ward. They must also include complications from all EOC facilities in the area under study in the numerator.

UNICEF/WHO/UNFPA has set the minimum acceptable level of "met need" as 100 percent, but in most developing country settings, this target is unrealistic. If evaluators find less than 100 percent, they conclude that some women with complications are not receiving the necessary medical care. However, if "met need" is low, researchers should seek other data to determine whether the problem lies in the availability, accessibility, qual-

---

[10] Much of the text for this indicator comes from "Indicators for Reproductive Health Program Evaluation," The EVALUATION Project, (Bertrand and Tsui, 1995), originally based on *Guidelines for Monitoring Progress in the Reduction of Maternal Mortality*, (Maine, McCarthy and Ward, 1992), and the 1997 version *Guidelines for Monitoring the Availability and Use of Obstetric Service*s (Maine et.al, 1997).

ity of care provided, or other factors, such as cultural factors, that determine the utilization of services.

Theoretically "met need" can exceed 100 percent, if more than 15 percent of pregnant women in the population develop major obstetric complications. In developed countries, the proportion of women with complications managed in EOC facilities may be greater than 15 percent of all births. Over-diagnosis of complications, which is seen in parts of Eastern Europe, can also cause this ratio to exceed 100 percent.

One difficulty with "met need" is that complications are subject to numerous recording biases and, even when standard definitions are in place, results can vary greatly with the data collection system being used and the training of the staff.

"Met need" is also particularly sensitive to the number of abortions included in the numerator. If the incidence of unsafe abortion is high, "met need" is likely to be high. The inclusion of all abortions can cause "met need" to be twice or three times as high as it would be without the abortions. Given this inflation of "met need" as a result of the inclusion of all abortion complications, a growing number of advocates for the indicator calculate it both ways, with and without all abortions. By excluding postabortion complications, estimates may be more comparable.

The appropriateness of using 15 percent of all births/pregnancies to estimate the number of women who experience obstetric complications is also open to discus-

sion. WHO's estimates of births with complications may be higher than 15 percent: hemorrhage, 10 percent of pregnancies; sepsis, 8 percent; hypertensive disorders of pregnancy, 5 percent; obstructed labor, 5 percent (WHO, 1996a). However, prospective data from West Africa suggest that 6 percent more reasonably estimates severe obstetric complications (Prual, 2000). The narrower the definition of what is considered a direct or major obstetric complication, the more reliable and comparable the estimates will be (MotherCare, 2000a). However, birth records and registries will likely lack sufficient detail on complications to allow much refinement regarding the severity of a complication.

The issue of double-counting a woman in the numerator (one who is admitted to the same facility more than once during her pregnancy or postpartum period or one who is admitted to more than one facility) is unlikely to seriously bias the results. If this situation were to occur, it would bias the indicators by presenting a more positive view of the health system than merited.

Given that the crude birth rate (CBR), the total population, and 15 percent are all estimates and that the accuracy of the CBR and population may vary according to the source, "met need" will likely be imprecise and may over- or underestimate the true value. To make the indicator useful for comparisons across facilities and districts or over time, one must use the same definitions and document the criteria used in each definition.

Part III.E
Newborn Health

- Proportion of hospitals and maternity facilities that are designated as "Baby-Friendly"
- Percent of audience who know about the warning/danger signs of newborn complications
- Percent of pregnant women with at least two doses of tetanus toxoid immunization
- Percent of home births with clean cord care
- Percent of newborns attended during the postnatal period by a health care provider
- Percent of live births with low birth weight
- Number of neonatal tetanus cases
- Neonatal mortality rate (NMR)
- Perinatal mortality rate (PMR)
- Birth weight specific mortality rate (BWSMR)

The goal of many programs in developing countries is to improve maternal and newborn health and survival. Until recently, however, newborn health has been relatively neglected in both the international child and safe motherhood movements, and few programs have focused specifically on improving newborn survival. A prime reason that newborn health has received such low priority is the general lack of awareness of the sheer numbers of early infant deaths. WHO estimates that each year more than 8 million infants die in the first year; of these, almost two thirds (5 .1 million) die in the first month, and of these, two thirds die within the first day (McCarthy, Lawn, and Ross, 2001). Virtually all of these deaths occur in developing countries. Although post-neonatal mortality has declined substantially, neonatal deaths have declined slightly and hence represent a growing proportion of the overall number of infant deaths (Espeut, 1998).

A second factor has been the perception that sophisticated technologies are required to significantly reduce newborn mortality. On the contrary, most newborn deaths in developing countries can be prevented by interventions already widely used. The most common causes of mortality – infections, asphyxia and birth injuries – can be prevented by simple cost-effective interventions that also benefit the mother. These interventions include antenatal malaria prevention and treatment, tetanus toxoid immunization, the detection and management of sexually transmitted infections, and access to a clean and safe delivery (WHO, 1996a). Furthermore, providing all infants with an "essential package of newborn care" (see Table III.E.1) including appropriate resuscitation, warmth, cleanliness and hygiene, clean cord care, and early exclusive breastfeeding also increases survival and reduces the proportions of surviving infants with disability (WHO, 1996b; WHO, 2001a).

Compared to other programmatic areas in Part III of this *Compendium*, newborn health is one of the least developed. Operations research studies have identified which interventions are likely to effectively reduce newborn mortality, but how these services should be scaled up, by whom, and at what cost must still be determined. The evaluation of these programs is therefore also in its infancy, and many new data-gathering tools, analytical approaches, and indicators need to be developed and tested. Because of the close link between maternal and newborn health, however, many process indicators appropriate for newborn health have already been used extensively in safe motherhood programs. (See Figure III.E.1) Indeed, separating maternal and newborn health indicators into two distinct parts may appear a false dichotomy when the antecedents of a poor pregnancy outcome, and the program interventions required to address these, may be the same for both mothers and babies. However, our purpose in doing so is to acknowledge growing awareness of the importance of newborn health and to highlight the fact that despite the many parallels between maternal and newborn health programs, important differences influence the way that programs are monitored and evaluated. These differences arise, not only because program interventions may vary, but because interventions that benefit both mothers and babies may differentially affect mortality. For example, because of their greater impact on newborn survival, interventions such as immunizing pregnant mothers against tetanus and detecting and treating STIs are more likely to be monitored in newborn health programs than in safe motherhood programs.

Methodological Challenges of Evaluating Newborn Health Programs

Some of the challenges of evaluating newborn health programs are similar to those for safe motherhood: the need to consider two outcomes, the large number of proximate determinants, and the difficulties of attributing causality to certain interventions because services are "bundled." Even though newborn deaths are more frequent than maternal deaths and therefore easier to count, the several measurement challenges include, but are not limited to, the following:

- **Countries define births, deaths, and "newborn period" in different ways, making valid international comparisons difficult.**

Meaningful use of any indicator is only feasible when standard definitions are used and applied. The first challenge for managers of newborn health programs is the lack of a generally agreed-upon definition of the "newborn period." In some settings "newborn" may refer to infants up to a few days of age and in other settings to infants up to several weeks of age. In this *Compendium,* the term newborn refers to the neonatal period (i.e., the first 27 completed days of life).

A second challenge is the different ways of aggregating newborn deaths according to the timing of the death. Typically, deaths are aggregated in the first month or first week of life or as fetal deaths (stillbirths and reported as described in Table III.E.2).

Many countries, however, define and record births and deaths in ways that may differ from the standard definitions of fetal, perinatal, and neonatal deaths recommended by the WHO in ICD 10. (See Table III.E.2). Some countries, for example, only record a baby as a live birth if the baby survives beyond 24 hours (McCarthy, Lawn, and Ross, 2001). The definition of a perinatal death is particularly problematic, not only because legal reporting requirements may vary between countries, but because the two standard international definitions have different gestational age and weight criteria. (See indicator **Perinatal Mortality Rate**.)

Further difficulties may also arise because national birth and death criteria may be differently interpreted and applied in different settings, and thus live births may be misclassified as fetal deaths and vice versa. This problem may occur because of an individual's lack of training or experience (this is true of deaths that occur in facilities as well as at home) or because of the way institutions and public health authorities choose to interpret and apply national birth and death criteria. Changes in medical practice may also alter the way practitioners systematically classify deaths, and managers need to be aware of this situation when they introduce programs to improve the quality of newborn health services at a facility level.

- **Ideally, all deaths (including fetal, perinatal, and neonatal) should be counted, but in practice, counting neonatal deaths is often the only feasible approach.**

In any program, the types of deaths to be counted depend on the program objective, the measurement method chosen, and the resources available. Ideally, program staff should record information on all neonatal, perinatal, and fetal deaths in order to derive a complete picture of pregnancy outcome. Having a complete picture is important because the causes of stillbirths and neonatal deaths are often the same, and the distinction between a stillbirth and a neonatal death may in practice be a very fine one. Just examining one rate or the other may underestimate the true level of mortality in the newborn period. Counting all deaths will provide a measure of the impact of interventions on etiologies that cause both fetal and neonatal deaths and will reveal the changes in the epidemiology of newborn deaths. Modern technologies, for example, may improve delivery outcome but increase neonatal mortality because fetal deaths are displaced from the antepartum to the post partum period. Reporting neonatal, perinatal, and stillbirth rates together will also permit more valid comparisons across programs and settings. Realistically, however, few programs will be able to achieve this goal. In most developing countries, the majority of births and deaths occur at home (WHO, 1996b). Few countries have sufficiently well-developed vital registration systems that can provide valid and reliable information on all births and deaths in the community. Health information systems can only provide information on facility births and deaths and, in most settings, are also poorly developed. Most community-based programs, if they have the capacity to measure mortality at all, will generally only be able to collect valid data on neonatal deaths for reasons explained below.

- **The quality of newborn mortality data is poorer than the quality of data for other ages.**

Evaluators require information on all births **and** deaths so that they may derive valid measures of newborn health outcome. As mentioned above, few developing countries have sufficiently developed vital registration systems to provide this data and, in many settings, re-

porting is very incomplete. Institutional barriers leading to underreporting may include cost, distance to the registration office, and a lack of awareness of the importance of birth and death registration. Social and cultural barriers may also make it difficult to collect valid data.

There are also major biases in the way deaths are reported. Even in countries with well-developed registration systems, a bias exists towards the reporting of larger, older babies, whereas deaths of very small babies early in the neonatal period are often omitted. Fetal deaths are much less likely to be reported than deaths of live births (WHO, 1996b).

- **Survey estimates of newborn mortality may not be suitable for short-term monitoring.**

Prospective studies would provide the most reliable mortality rates but are too expensive for regular reporting purposes. In practice, the most reliable estimates of neonatal mortality are derived from large scale surveys that rely on the retrospective report of deaths in early infancy. Surveys focusing on live births provide estimates of neonatal mortality, but perinatal mortality estimates require complete pregnancy histories. Because many population-based surveys focus on obtaining demographic indicators that use live births in the denominator, there has been relatively less experience with the use of pregnancy histories (which collect information about stillbirths). The reliability of any survey estimate depends on the completeness of reporting, and underreporting is generally more pronounced for deaths in early infancy. Because of the relatively small numbers of deaths recorded in this type of survey, national neonatal mortality rates are usually presented for a period of five years before the survey, and sub-national estimates are presented for ten years before the survey. The lack of precision in the estimates may sometimes make it very difficult to assess the significance of small changes between surveys (Rutstein, 1999).

- **Measuring perinatal and neonatal morbidity is very difficult.**

Estimates of newborn morbidity are important for designing effective program interventions. As with safe motherhood, however, existing estimates of newborn morbidity are usually derived from facility data and are unlikely to reflect the true burden of morbidity in the community unless all births and deaths are institutional.

Although community members can learn to diagnose illness in a sick newborn (Bang et al., 1999), illness is often difficult to recognize because babies usually present with relatively non-specific symptoms, such as poor feeding and lethargy. Assigning a cause of death may be difficult because many different diseases may present with the same symptoms, and many babies die at home before ever reaching medical attention. Few facilities have adequate diagnostic facilities when ill babies do eventually present for care.

- **New program indicators are required at the individual, community, and facility level.**

Much of the discussion on the challenges of monitoring and evaluating newborn health has so far focused on newborn mortality because of the relative lack of experience with process indicators for newborn health. Although mortality indicators clearly have their place and provide the only direct measure of the objective of most programs, process indicators need to be developed to measure the wide range of interventions required to improve newborn health and survival.

Process indicators are required for measuring the availability, accessibility, quality, and demand for services at the facility level where the provision of newborn health services has historically been overlooked. A recent national survey in Kenya, for example, showed that over one third of hospitals lacked even the most basic equipment for resuscitation (MOH, NCPD, and ORC Macro, 2000). (See Part II.H.2a Service Provision Assessment on quality of newborn care services and service delivery.)

In addition to monitoring at the facility level, indicators are also required for monitoring and evaluating interventions at the individual and community level. Many infants become ill and die before ever reaching medical care. It is particularly important to develop indicators that help programs understand community knowledge, attitudes and behaviors in response to newborn illness and to determine which interventions are the most effective.

The Maternal and Newborn Health Framework

The maternal and newborn health conceptual framework (see Figure III.E.1) from which many indicators in this manual were developed illustrates the links between maternal and newborn health from before pregnancy to

after delivery. The framework also shows where interventions can promote and improve health status as well as reveal the levels (family, community, and services) at which the impact of these interventions should be measured. The framework does not address some of the system-level determinants – the social, cultural, economic, political, and legal factors that clearly also influence maternal and newborn health. The indicators included in this section of the *Compendium* relate directly to newborn health and newborn health care. Many other indicators that also affect newborn health such as birth timing (**Contraceptive Prevalence Rate**) and nutrition (**Percent of Women of Reproductive Age with Anemia**), appear in other sections of the *Compendium*.

The Selection of Indicators

Most indicators in this section of the *Compendium* are intended for use at the national level or in the context of large-scale programs, but many can be used in other contexts. Using the same criteria applied to the safe motherhood selection, a small expert group currently working in newborn health programs in wide consultation with other experts and organizations working in newborn health selected these indicators.

The indicators measure predominately health outcomes and impacts because these are the newborn health program indicators most widely used to date. We are aware that some programs cannot measure these outcomes, and for these programs, our selection will be less useful. However, in contrast to safe motherhood, some programs, even at an NGO level, may encounter sufficient numbers of deaths to derive stable estimates of newborn mortality. Moreover, these programs are particularly well placed to contribute much-needed research on the relationship between community behaviors and newborn health outcomes.

The relative absence of newborn health process indicators reflects the very recent development of newborn health as a distinct programmatic area. An urgent need clearly exists for research to develop our understanding of how best to monitor and evaluate newborn health programs, but until further work validates proposed indicators, many can only be termed "experimental" and do not meet this *Compendium's* criteria. Nevertheless, we include a small number of indicators that have been field tested by certain groups but not yet widely adopted. They are the **Percent of Home Births with Clean Cord Care** and the **Percent of Audience who Know about the Warning/Danger Signs of Newborn Complications**.

These indicators are included because of the need to stimulate debate and discussion on appropriate newborn health indicators, even though we recognize that these two indicators may not meet all the criteria for a good indicator (WHO, 1997) and that neither are particularly intended for national level use. In addition, the decision to include the **Percent of Newborns Attended during the Postnatal Period by a Health Care Provider** provoked considerable debate because of the lack of consensus on the timing and objectives of postnatal care. We retain the indicator because it is in widespread use, and one objective of the *Compendium* is to promote standardization of definitions and concepts.

In the next few years, as awareness of the problem of newborn mortality grows, no doubt those working in newborn health will move toward consensus on the indicators appropriate for monitoring national level programs. This set of indicators represents one small step in that direction.

Figure III.E.1 MATERNAL & NEWBORN HEALTH FRAMEWORK
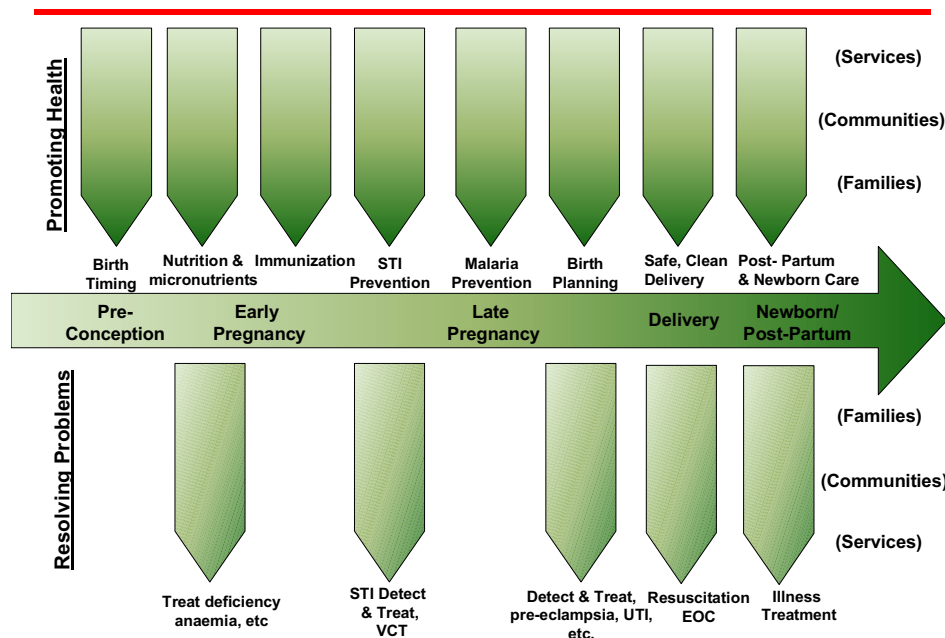
Adapted from Al Bartlett, USAID.

Table III.E.1

WHO Essential Newborn Care Package

1. Cleanliness: clean delivery and clean cord care for the prevention of newborn infections (tetanus and sepsis)

2. Thermal protection: prevention and/or management of neonatal hypo-thermia and hyperthermia

3. Early and exclusive breastfeeding

4. Initiation of breathing, resuscitation

5. Eye care: prevention and management of ophthalmia neonatorum

6. Immunization (BCG, Oral polio, Hepatitis B)

7. Management of newborn illness

8. Care of the preterm and/or low birth weight newborn

Source: WHO (1996b)

Table III.E.2
Standards and Reporting Requirements Related
to Perinatal and Neonatal Mortality

Stillbirth

A stillbirth is the death of a fetus weighing 500g or more, or of 22-weeks gestation or more, if weight is unavailable. Because gestation and birth weight are often unavailable for early fetal losses, for international comparisons, the WHO recommends including only deaths of fetuses weighing at least 1000g, (or of 28-weeks gestation or more if weight is unavailable).

The terms stillbirth and fetal death are sometimes used interchangeably. This text chooses the term stillbirth because it is most widely recognized.

Neonatal Mortality Rate (NMR)

$$\frac{\text{\# of neonatal deaths (deaths of live births within the first 28 completed days of life)}}{\text{\# of live births}} \times 1000$$

Early Neonatal Mortality Rate (ENMR)

$$\frac{\text{\# of early neonatal deaths (deaths within the first 7 completed days of life)}}{\text{\# of live births}} \times 1000$$

Late Neonatal Mortality Rate (LNMR)

$$\frac{\text{\# of late neonatal deaths (deaths within 7-27 completed days of life)}}{\text{\# of live births}} \times 1000$$

Perinatal Mortality Rate (PMR)

$$\frac{\text{\# of stillbirths + \# of early neonatal deaths}}{\text{Total \# of births (stillbirths + live births)}} \times 1000$$

Note: The day of birth is counted as Day 0, so that "within 1 week" includes babies 0-6 days old and "within 1 month" includes babies 0-27 days old.

Source: ICD10 (WHO, 1992)

PROPORTION OF HOSPITALS AND MATERNITY FACILITIES THAT ARE DESIGNATED AS "BABY FRIENDLY"

Definition
The proportion of hospitals and maternity facilities that have been accredited as "Baby Friendly" according to the ten UNICEF/WHO criteria related to breastfeeding and newborn care

To be designated as "Baby Friendly," the hospital must:

1. Have a written breastfeeding policy that is routinely communicated to all health care staff;
2. Train all health care staff in skills necessary to implement this policy;
3. Inform all pregnant women about the benefits and management of breastfeeding;
4. Help mothers initiate breastfeeding within an hour of birth;
5. Show mothers how to breastfeed and how to maintain lactation, even if they should be separated from their infants;
6. Give newborn infants no food or drink other than breast milk, unless medically indicated;
7. Practice "rooming in" by allowing mothers and infants to remain together 24 hours a day;
8. Encourage breastfeeding on demand;
9. Give no artificial teats, pacifiers, dummies, or soothers to breastfeeding infants; and
10. Foster the establishment of breastfeeding support groups and refer mothers to them on discharge from the hospital or birthing center.

This indicator is calculated as:

> # of hospitals and maternity facilities accredited as "Baby Friendly"
> _____
> The total number of maternities and hospitals that handle deliveries

Data Requirements
The number of maternities meeting BFHI criteria; the total number of maternities and hospitals

Data Source(s)
UNICEF/WHO/Wellstart Baby Friendly Hospitals Initiative internal self-assessment and external evaluation instruments

Purpose and Issues
The Baby Friendly Hospitals Initiative (BFHI) is a joint UNICEF/WHO/Wellstart initiative aimed at increasing breast-feeding rates and encouraging global standards for maternity services in hospitals and maternities. Facilities first conduct a self-assessment; then independent assessors appointed by the national BFHI committee or UNICEF country offices evaluate them according to the above criteria. These same bodies aggregate information on the numbers and proportions of facilities acquiring "Baby Friendly" status for national and global reporting (WHO, UNICEF, and Wellstart International, 1999).

Whereas this indicator provides useful information on the availability of baby-friendly services in a given country, we cite several caveats in its use. First, the number of facilities achieving "Baby Friendly" status may be presented more often than the proportion because of difficulties in ascertaining the total number of maternities required for the denominator. The number of facilities is clearly affected by country size. For example, by December 2000, 6312 hospitals in China (or 47 percent of all eligible facilities) had achieved "Baby Friendly" status compared to 232 (or 66 percent of all eligible facilities) in Kenya. Ascertaining the number of maternities in the private sector is particularly difficult, and in many cases, private facilities may not be represented in national estimates.

Second, the listing of facilities recorded as "Baby Friendly" may be out of date because periodic reaccredidation to maintain standards is voluntary and depends on the interest and motivation of each individual facility. The date of acquiring "Baby Friendly" status and whether reaccredidation has occurred are not routinely recorded.

Percent of Audience Who Know about the Warning/ Danger Signs of Newborn Complications

Definition

Community knowledge and awareness of the warning/ danger signs of newborn complications

The "audience," the intended population for the program, will usually be mothers in the case of newborn babies. Husbands or other household members known to influence decisions about care seeking, as well as other health care providers (such as traditional birth attendants), may also need to know about signs of newborn illness.

"Know" refers to the percentage who can spontaneously name the warning/danger signs of newborn complications, (for example, the percentage of mothers who can name at least three commonly recognized signs of newborn illness).

This indicator is calculated as:

$$\frac{\text{\# respondents who know the warning/danger signs of newborn complications}}{\text{Total \# of respondents}} \times 100$$

Data Requirements

Response to knowledge questions asked in surveys

Data Source(s)

Population-based survey, preferably with a representative sample of the audience

Purpose of Issues

The purpose of this indicator is to assess community knowledge and awareness of the warning/danger signs of newborn complications in order to plan and monitor the impact of BCC program efforts at a community level.

Because most babies are born at home or are discharged from the hospital in the first 24 hours, increasing community awareness of the danger signs of newborn complications is of critical importance for improving new-born survival. More babies die in the first week of life than at any other time in childhood, and those who become ill shortly after birth may deteriorate and die very rapidly. The warning signs of newborn illness may not be recognized, because they are often much less pronounced than those in an older child or adult. Community members can, nevertheless, learn to recognize signs and symptoms of newborn illness (Bang et al., 1999).

The limitations of assessing community knowledge of signs and symptoms of newborn illness are similar to those outlined for obstetric complications. (See indicator **Percent of the Audience Who Know About the Danger/Warning Signs of Obstetric Complications.**)

A major limitation with newborn complications is that little consensus exists on which signs and symptoms the general public can use to improve the early diagnosis of serious illness at the community level. Algorithms shown to be sensitive and specific in clinical settings are too complex for use by the general public (McCarthy, Lawn, and Ross, 2001). More simple measures are less specific and will lead to larger numbers of newborns receiving unnecessary treatment. However, having some healthy babies over-treated is preferable to having some sick babies being under-treated and dying as a result. Danger signs that have been proposed include:

- Breathing difficulty, irregular or fast (>60 minute);
- Feeding poorly (less than half of usual consumption);
- Jaundice, pallor, bleeding;
- Convulsions, spasms, jitters;
- Fever temperature greater that 38°C or low temperature less than 36°C; and
- Vomiting green, no stool in 24 hours of life, swollen abdomen.

(McCarthy, Lawn, and Ross, 2001).

Programs aimed at raising community awareness of neonatal illness should carry out formative research to determine what signs of illness are already recognized in

the community and how to adapt general recommendations to a specific setting. More fundamental research is required to reach consensus on which signs and symptoms caretakers in different settings can consistently recognize.

PERCENT OF PREGNANT WOMEN WITH AT LEAST TWO DOSES OF TETANUS TOXOID IMMUNIZATION

Definition

The proportion of pregnant women receiving at least two doses of tetanus-toxoid vaccine (TT2)

This indicator is calculated as:

$$\frac{\text{Total TT2 + TT3 + TT4 + TT5}}{\text{Total \# of live births}} \times 100$$

Where TT2, TT3, TT4, TT5 refer to the 2nd, 3rd, 4th, or 5th dose of tetanus-toxoid vaccine administered (WHO, 1999a and c).

Data Requirements

From service statistics:
Number of doses of TT2 + TT3 + TT4 + TT5 given to pregnant women in a reference period (usually a year)

From population-based surveys:
Number of women giving birth during a reference period (e.g., five years) who report receiving at least two doses of tetanus-toxoid during their last pregnancy and number of live births in the same reference period

The number of live births serves as a proxy for the number of pregnant women.

Where data on the numbers of live births for the denominator are unavailable, evaluators can calculate total estimated live births using census data for the total population and crude birth rates in a specified area. *Total expected births = population x crude birth rate*

In settings where the crude birth rate is unknown, the WHO recommends using 3.5 percent of the total population as an estimate of the number of pregnant women (number of live births or pregnant women = total population x 0.035) [WHO, 1999a and c.]

Data Source(s)

Service statistics; population-based surveys

Purpose and Issues

This indicator measures the percentage of women and births protected against tetanus at the time of delivery among clients in a given program or among the general population.

Neonatal tetanus is usually fatal. A woman immunized with at least two doses of tetanus toxoid according to the WHO schedule[1] develops antibodies that protect her infant against tetanus in the first two months of life. Tetanus-toxoid immunization is therefore an integral part of the ANC package offered to women in most developing countries.

Many national HIS routinely collect this indicator to provide TT2+ coverage estimates for women attending facilities for ANC. Most large population-based surveys also collect data on self reported TT2+ coverage.

Note: Variations in the methods used to measure TT2+ coverage, as well as in the definition of the numerator and denominator, give rise to differences in the magnitude and reliability of the estimates obtained. For example, service statistics record the total number of doses of a vaccine in the previous *12 months*, whereas surveys tend to record the total number of women who report receiving at least two vaccinations during their *last pregnancy in a reference period that may be up to five years.*

Service statistics have the disadvantage that they may be incomplete or inaccurate (WHO, 1999a). They are also subject to a selection bias and are not representative of the general population, particularly when ANC coverage is low. However, they provide the only way of monitoring coverage on an annual basis and may be more reliable than self-reported data are.

---

[1] Given at least four weeks apart with the second dose administered before the 36th week of pregnancy.

Surveys provide the only means of obtaining population-based coverage, but because surveys rely on self-reporting, they are subject to recall bias that is likely to increase with the length of the recall period.

Both approaches, however, underestimate the true extent of TT2+ coverage because both exclude doses of vaccine administered at times other than specified in the definition of the numerator even though the doses offer protection. For example, the doses for the childhood or mass-immunization campaign are omitted.

Promoting clean delivery and cord care practices as well as ensuring that women are adequately immunized against tetanus prior to birth can prevent transmission of neonatal tetanus. TT2+ coverage should also be reported as well as the number of neonatal tetanus cases and the proportion of live births with a skilled attendant (as a proxy for clean births).

For prevention of neonatal and maternal tetanus, WHO recommends giving women a series of five doses of tetanus-toxoid vaccine with a minimum interval between each dose. Each dose increases the level and protection against tetanus. Each dose counts as a dose towards a five-dose schedule even if given before the recommended interval. A woman who receives five doses of tetanus toxoid is fully immunized and is protected against tetanus throughout her childbearing years.

Table III.E.3 WHO Recommended Tetanus-Toxoid Series			
TT Dose	**Given**	**Level of Protection**	**Duration of Protection**
TTI	At first contact	NIL	None
TT2	Four weeks after TTI	80%	3 years
TT3	At least 6 months after TT2	95%	5 years
TT4	At least one year after TT3	99%	10 years
TT5	At least one year after TT4	99%	30 years

PERCENT OF HOME BIRTHS WITH CLEAN CORD CARE

Definition

The coverage of clean cord care at the time of home delivery, either from use of a clean delivery kit or a new blade to cut the cord (KPC, 2000)[2]

This indicator is calculated as:

$$\frac{\text{\# of home births with a clean delivery kit or clean blade}}{\text{Total \# of home births}} \times 100$$

Clean births kits typically include at least soap, a new razor blade, cord ties, and a plastic sheet.

Data Requirements

Number of home births with a clean delivery kit or clean blade in a defined geographical place and time, based on self-reports of women with children 0-23 months and number of home births in the same place and time period

Where data on the numbers of live births for the denominator are unavailable, evaluators can calculate total estimated live births from the total population and crude birth rate in a specified area. *Total expected births = population x crude birth rate.*

This indicator is most appropriate in settings where facility births are rare. If facility births occur in significant numbers, then evaluators should adjust the denominator accordingly.

Data Source(s)

Population-based survey (national, regional, district)

Purpose and Issues

This indicator measures the coverage of clean-delivery and cord-care practices at birth for deliveries that take place outside a facility.

Tetanus and sepsis are two leading causes of maternal and neonatal morbidity in the developing world. These deaths result from contamination from an unclean environment but could be prevented with improved hygiene and cord care at the time of delivery. Use of clean home delivery kits and a new blade for cutting the umbilical cord have been shown to reduce the incidence of simple cord infection, but no studies have assessed the impact on mortality or more serious infections because of the need for very large sample sizes (Tsu, 2000). Clean cord care is one of the key elements in the Essential Newborn Care Package (WHO, 1996a).

Two caveats warrant mention. First, surveys that rely on a women's recall of events at the time of delivery are subject to a recall bias likely to increase with the length of the recall period. Furthermore, if a woman was attended by a TBA, she may be unaware whether a clean delivery kit was used or how the cord was cut. A courtesy bias may affect the response if respondents are aware that programs are known to be promoting certain delivery practices.

The many elements to clean cord care and a number of alternative or complementary indicators include the following:

- Percent of births in which cord was cut with a new blade on a clean surface;
- Percent of the intended audience who know the importance of clean cord care;
- Percent of the intended audience who intend to observe clean cord care at their next birth; and
- Percent of the intended audience who express satisfaction with the use of clean cord care at their last delivery (Koblinsky et al., 1995; Tsu, 2000).

In addition to including beneficial practices, some programs may wish to monitor the reduction in potentially harmful practices that encourage the spread of tetanus. For example:

- Percent of babies whose cord stump was treated with dung or ashes.

---

[2] Available online at: http://www.childsurvival.com/kpc2000/kpc2000.cfm

PERCENT OF NEWBORNS ATTENDED DURING THE POSTNATAL PERIOD BY A HEALTH CARE PROVIDER

Definition
The percent of newborns attended by a health care provider during the postnatal period

This indicator is calculated as:

$$\frac{\text{\# of newborns attended during the postnatal period by a health care provider}}{\text{Total \# of live births}} \times 100$$

The postnatal period begins one hour after the birth of the placenta and ends 6 weeks later (WHO, 2001b). Although not officially defined, for the purposes of this indicator, the postnatal period can be roughly divided into 3 stages: (1) immediate postnatal period (first 2 hours after birth), (2) early postnatal period (from 2 hours to 12 hours after birth), and (3) late postnatal period (from day 2 to day 40).

Data Requirements
Numbers of newborns who are attended during the postnatal period (the numerator should specify whether newborns are seen for the first or subsequent visit); all live births in the same period

The indicator should be calculated specifying the stage of postnatal period (for example: percent of newborns attended within the first two days of birth).

Data Source(s)
Service statistics; population-based surveys

Routine HIS may collect data for this indicator to obtain estimates of postnatal coverage. Routine health service data generally lacks information on pregnancies or births that take place outside the public sector, for example in homes or in private sector facilities, and therefore should serve to estimate a denominator.

Where data on the numbers of live births for the denominator are unavailable, evaluators can calculate total estimated live births using census data for the total population and crude birth rates in a specified area. *Total expected births = population x crude birth rate*

Purpose and Issues
The main purpose of an indicator for postnatal care is to provide information on the use of postnatal services and to provide a measure of access to services for newborns in the postnatal period.

More deaths occur in the early postnatal period than at any other time of life. Relatively little attention has focused on developing postnatal services, however, and no clear recommendations on the optimal frequency, timing, content and delivery of postnatal care exist. Consensus is beginning to emerge in a number of areas. For example:

- Ideally, all women and newborns should be seen together in the postnatal/partum period to permit the early detection and treatment of complications that occur commonly after birth and to provide preventive care to both mother and baby (WHO, 1998a; WHO, 2001a).
- Preventive care for mothers should include vaccination against tetanus; provision of vitamin A and iron; and counseling on appropriate newborn care, hygiene, breastfeeding, malaria prevention and nutrition. Preventive care for babies should include early BCG, polio, and hepatitis immunization (WHO, 1998b; WHO, 2001a); and
- In the absence of problems, postnatal visits should take place at 6-12 hours, 3-7 days, and 4-6 weeks. The major priority is for a visit in the first 24 hours. (WHO, 1998a).

Many routine HIS and population-based surveys collect data on postnatal care coverage. Annual statistics are only possible through HIS. Surveys of PNC coverage every three to five years are sufficient. More frequent measurement is inappropriate because sampling errors make it difficult to assess whether small changes are real or due to chance variation.

Postnatal care coverage should respond to program interventions aimed at increasing coverage in the short term.

Several additional points concerning the interpretation of this indicator are worth emphasizing.

First, the lack of an agreed-upon operational definition of postnatal care makes valid international comparisons difficult. Postnatal care is a package of services and not one single intervention. Because the content and quality of care are likely to vary between settings, similar coverage rates do not necessarily reflect similar levels of care.

Second, postnatal coverage rates should make explicit whether care was provided principally for the mother or baby, or both mother and baby, because this detail may be difficult to determine retrospectively. The current DHS questionnaire, for example, asks about postpartum care for the mother, but not for the baby, and routine HIS may not make such distinctions.

Third, postnatal care coverage should ideally be stratified by the age of the baby after birth to get a better measure of access to services in the immediate postnatal period. Routine HIS usually collects data on postnatal coverage without specifying when this visit took place.

Finally, surveys relying on a woman's recall of events are subject to a recall bias likely to increase with the length of the recall period.

PERCENT OF LIVE BIRTHS WITH LOW BIRTH WEIGHT

Definition
Low birth weight (LBW), which is defined as a body weight at birth of less than 2,500 grams (g)

This indicator is calculated as:

$$\frac{\text{\# of births} <2500g}{\text{Total \# of live births}} \times 100$$

LBW has two main causes: preterm birth and intrauterine growth retardation (IUGR). LBW is often used as a proxy indicator to quantify the magnitude of IUGR in developing countries because valid assessment of gestational age is generally not available.

Preterm birth: The term preterm birth is used for infants born before 37-weeks completed gestation. Most, but not all, premature newborns in developing countries weigh less than 2500g.

Intrauterine growth retardation (IUGR): a condition in which fetal growth has been impaired. In developing countries, maternal under-nutrition and maternal ill health including malaria, anemia and acute and chronic infections (e.g., STIs) are major causes.

Data Requirements
Number of newborns with a birth weight less than 2,500g in a defined time period (e.g., 12 months) and number of live births in the same time period

Data Source(s)
Population-based surveys; health services data

Routine HIS may collect data for this indicator to obtain estimates of LBW for facility births.

Purpose and Issues
Approximately 1 in 6 newborns, or 17 million babies, are born every year with low birth weight. Low birth weight is the single most important predictor of newborn well-being and survival. Because maternal under-nutrition is a major determinant of LBW, high rates of LBW should be interpreted not only as an indicator of newborn under-nutrition, morbidity, and mortality, but also as an indicator of maternal well being. One of the goals of the World Summit for Children is to reduce the incidence of low birth weight to less than ten percent (ACC/SCN, 2000a).

In developing countries, approximately two thirds of LBW is caused by IUGR, and the remaining one third is due to preterm birth, although some preterm babies also have IUGR. By contrast, in developed countries, the majority of low birth weight is due to preterm birth.

Low-birth-weight babies are ten times more likely to die than babies weighing over 3 kg. They are also more likely to have impaired cognitive development and to develop acute illnesses such as diarrhea and pneumonia in early infancy (ACC/SCN, 2000a).

Obtaining reliable estimates of low birth weight in the general population is difficult. In many developing countries, the majority of births occur at home and babies are not weighed; thus, the data that are available come from a relatively small proportion of facility births.

Many household surveys collect data on birth weight, but since the weights reported are mainly from facility births, these data are also subject to selection bias. Some household surveys (such as the DHS) ask mothers to state whether their baby was smaller than average or very small; and at an aggregate level these data may be used to estimate incidence of low birth weight at a national level. Regional estimates are also possible if the sample size is sufficiently large (Boerma et al., 1996).

This indicator measures one of the major objectives of safe pregnancy/neonatal interventions: to prevent low birth weight. However, since low birth weight is due to many complex factors, changes in low-birth-weight incidence occur slowly. Estimates every five years are probably reasonable and consistent with the schedules of many large surveys (e.g., the DHS). Evaluators must

recognize that this indicator will be slow to change, even with well-executed interventions.

Several caveats pertain to LBW. First, aggregate figures of low-birth-weight incidence may hide important differentials between sub-groups at risk. Second, heaping of birth weight recording in multiples of 500g is common and affects the incidence of low birth weight. Heaping is particularly a problem with survey data but also affects facility data to some degree.[3] Third, survey data rely on women's reports of their infant's birth weight and are subject to recall bias. Validation studies from the United States suggest that mothers are able to recall their baby's weight accurately, but we are not aware of similar large-scale studies conducted in developing countries.

---

[3] Heaping occurs when respondents do not know the exact weight. Estimated weights are often reported on certain preferred weights, such as multiples of 100 or 500 grams.

NUMBER OF NEONATAL TETANUS CASES

Definition
The number of neonatal tetanus (NT) cases in a given year, in a defined population, including both suspected and confirmed cases

A suspected case: any neonatal death between 3-28 days of age in which the cause of death is unknown; or any neonate reported as having suffered from neonatal tetanus between 3-28 days of age and not investigated.

A confirmed case: any neonate with a normal ability to suck and cry during the first 2 days of life; and who between 3-28 days of age cannot suck normally and becomes stiff or has convulsions (i.e., jerking of the muscles) or both.

The basis for case classification is entirely clinical and does not depend on laboratory confirmation. NT cases reported from hospitals are considered confirmed (WHO, 1999a).

Data Requirements
Number of neonatal tetanus cases or deaths

Data Source(s)
Population based NT mortality surveys; neonatal tetanus surveillance systems; and population-based surveys (TT2+ coverage, number of live births)

Purpose and Issues
Neonatal tetanus is a major public health problem in the developing world. Each year approximately half a million infants and almost 50,000 mothers die from tetanus acquired around the time of delivery. Current efforts focus on eliminating NT by 2005 in the 57 countries still reporting the disease. (The Expanded Program of Immunization, EPI, defines elimination of tetanus as a reduction in the incidence to fewer than 1 case per 1000 live births in every district of every country) [WHO,1999a].

Because the case fatality is very high in most developing countries, the number of neonatal tetanus cases is often based on actual or estimated numbers of NT deaths.

In countries with tetanus toxoid immunization coverage (TT+) of over 90 percent and a clean delivery rate over 80 percent, the number of neonatal tetanus cases is taken as the number of neonatal tetanus deaths reported.

In countries with lower coverage, an estimate of the number of NT cases is based on an estimate of NT deaths calculated from the number of live births, the neonatal tetanus mortality rate (NTMR), TT2+ coverage, and vaccine efficacy (VE).

Some countries occasionally conduct NT mortality surveys, although most countries with a high proportion of neonatal tetanus deaths carry out routine surveillance in "high risk" areas. Unfortunately, surveillance systems function poorly, and neonatal tetanus continues to be seriously underreported. Community-based NT mortality surveys, for example, suggest that routine surveillance systems detect only two to eight percent of all cases (WHO, 1994b). For this reason, WHO recommends using the following calculation in most settings.

$$\text{\# of NT deaths in 1 year} = \text{Live births} \times \text{NTMR} \times (1 - \text{TT2+} \times \text{VE})$$

Where:

NTMR = the baseline Neonatal Tetanus Mortality Rate (mortality rate in unvaccinated cases);

TT2+ = Tetanus-toxoid-immunization coverage; and

VE = Vaccine efficacy (estimated as 0.95).

The NTMR used is the latest value reported in each country where a nationwide survey was undertaken; if no surveys were conducted, a rate of 1, 5, 10, 15 cases per 1000 live births is allocated on the basis of the NTMR reported in countries with similar risk factors. In Latin America the WHO Regional Office (AMRO) uses a correction factor for the sensitivity of the surveillance system to adjust for the numbers of reported neonatal tetanus deaths (WHO, 1994b).

Countries with NT surveillance systems assess their progress annually. Demographic surveys, providing neonatal mortality at 4-14 days on a 3-5 year basis, serve to evaluate surveillance data.

A number of caveats warrant mention. First, this indicator reflects the overall magnitude of the problem of neonatal tetanus deaths but does not offer a precise estimate because of serious underreporting from surveillance data and because of the many assumptions inherent in the WHO calculation. Second, because this indicator is reported as a number rather than as a propor-tion, countries with lower rates of NT deaths but larger populations will rank ahead of countries with proportionately higher deaths rates. Third, aggregate figures at a national level may disguise pockets of high risk in certain subgroups (for example in rural populations or low-caste groups).

Surveillance systems reporting the number of NT cases should also give the percent completeness of reporting (number of NT reports received/the number of reports expected in the same time period). Neonatal-tetanus deaths should also be reported in conjunction with TT2+ coverage and the proportion of live births with a skilled attendant (as a proxy for proportion of clean deliveries).

In countries where NT is a recognized problem, population-based surveys may provide information on levels and trends of neonatal mortality. These surveys provide information on neonatal mortality at 4-14 days, which is a sensitive indicator of NT mortality (Boerma et al., 1996).

Neonatal Mortality Rate (NMR)

Definition
The number of neonatal deaths per 1000 live births

A neonatal death is defined as a death during the first 28 days of life (0-27 days).

$$\frac{\text{\# of neonatal deaths}}{\text{Total \# of live births}} \times 1000$$

The NMR is often broken down into early and late mortality rates. The Early Neonatal Mortality rate (ENMR) is calculated as follows:

$$\frac{\text{\# of neonatal deaths 0-6 days}}{\text{Total \# of live births}} \times 1000$$

The late neonatal mortality rate (LNMR) is calculated as follows:

$$\frac{\text{\# of neonatal deaths 7-27 days}}{\text{Total \# of live births}} \times 1000$$

Data Requirements
Number of neonatal deaths in a given population and reference period and number of live births in the same population and reference period

Data Source(s)
Vital registration; population-based surveys; services statistics

Where data on the numbers of live births for the denominator are unavailable, evaluators can calculate total estimated live births using census data for the total population and crude birth rates in a specified area. *Total expected births = population x crude birth rate*

Routine HIS may collect data for this indicator to obtain estimates of the NMR for facilities. Facility data are not recommended for estimating the NMR for the general population, because in many settings, many neonatal deaths and live births occur outside the health system, which will cause substantial selection bias.

Purpose and Issues
The NMR is a key outcome indicator for newborn care and directly reflects prenatal, intrapartum, and neonatal care. In addition, as infant mortality rates decline, the proportion of infant deaths that occur in the neonatal period typically increases. The NMR differs from the perinatal mortality rate in that it focuses only on deaths among live births and covers a longer period after birth. Information on live births is generally thought to be easier to obtain than information on non-live births and is more widely available, because many population-based surveys such as the DHS typically only collect information on live births. Early neonatal deaths are more closely associated with pregnancy-related factors and maternal health, whereas late neonatal deaths are associated more with factors in the newborn's environment.

In many countries, vital registration data are not sufficiently complete to allow reliable estimation of the NMR. The standard techniques for collecting data on live births and neonatal deaths in population-based surveys have been widely applied in programs such as the WFS and DHS. Data quality is an important issue; common problems include omission of deaths, particularly very early neonatal deaths, and heaping of the reported age at death on 7, 28, or 30 days.[4] Heaping on these digits is particularly problematic because it will lead to the misclassification of early neonatal deaths as late neonatal death (7 days) or late neonatal deaths as post-neonatal deaths (28 and 30 days).

Evaluators typically calculate NMR at a national or international level. They may also obtain sub-national

---

[4] Heaping occurs when respondents do not know the exact age at death. Estimated ages at death are often reported on certain preferred ages, such as 7, 28, or 30 days, leading to a distorted age distribution of deaths in which too many deaths are reported at these preferred ages, and too few at the ages just before and after.

estimates if sample sizes are sufficiently large. The NMR is sometimes calculated at a facility level to monitor the outcome of delivery and newborn care in health facilities. Reliable estimates for individual facilities can only be obtained for very large facilities with large numbers of deliveries and neonatal admissions.

The NMR may respond fairly quickly to programmatic interventions, for example, immunizing all pregnant women in areas of high tetanus prevalence. However, survey-based estimates are generally subject to relatively large sampling errors, so it is impossible to detect changes over short periods of time unless the changes are quite large. Also, survey-based estimates are often based on a five-year period prior to the survey. Therefore, we recommend collecting survey-based estimates of the NMR not more than every three to five years.

One limitation of note is the NMR's sensitivity to changes in the quality of data. For example, a rise in the NMR may indicate deterioration in newborn health outcomes, or it may indicate an improvement in the reporting of neonatal deaths. Therefore, assessing data quality is essential to analysis.

Also, evaluators should interpret comparisons of facility-based estimates of the NMR very carefully because the NMR in a facility is very sensitive to the case mix of deliveries and neonatal admissions. One should not interpret a higher NMR in one facility as suggesting that the quality of neonatal care is worse in this facility because the NMR may rise or fall in response to changes in the case-mix. Additionally, improvements in prenatal and intrapartum care and advances in medical technology may increase the NMR because babies who may otherwise have been stillbirths may survive delivery only to die in the neonatal period. For these reasons, we recommend that evaluators break down facility-based estimates of the NMR by birth weight (see **Birth Weight Specific Mortality Rate**) and by admission status (direct admission or transfer-in) as a proxy for case mix.

PERINATAL MORTALITY RATE (PMR)

Definition
The number of perinatal deaths per 1000 total births

A perinatal death is a fetal death (stillbirth) or an early neonatal death.

The perinatal mortality rate is calculated as:

$$\frac{\text{\# of perinatal deaths}}{\text{Total \# births (still births+ live births)}} \times 1000$$

A stillbirth is the death of a fetus weighing 500g or more, or of 22-weeks gestation or more if weight is unavailable (ICD 10).

An early neonatal death (END) is the death of a live newborn in the first 7 days (i.e., 0-6 days) of life.

Great variation exists both between and within countries on how the stillbirth component of perinatal mortality is recorded, particularly for early stillbirths that occur at 22- to 27-weeks gestation. For international comparisons, WHO suggests including only deaths of fetuses weighing at least 1000g, or of 28-weeks gestation or more if weight is unavailable. Presentations of the PMR should include a clear statement of the definition of perinatal mortality used.

In practice, in most developing countries accurate data on birth weight or gestational age are difficult to obtain.

Data Requirements
Number of perinatal deaths in a given population in a given reference period (i.e., 12 months) and number of births (live births + stillbirths) in the same population and reference period

Data Source(s)
Population-based surveys; vital registration; service statistics

Routine HIS may collect data for this indicator to obtain estimates of the PMR for facilities. Facility data are not recommended for estimating the PMR for the general population because in many settings, many perinatal deaths and live births occur outside the health system, which will cause substantial selection bias.

Purpose and Issues
The PMR is a key outcome indicator for newborn care and directly reflects prenatal, intrapartum, and newborn care. It has also been proposed as a proxy measure of maternal health status and mortality, but a recent study has cast doubt on its use as a proxy for maternal mortality (Akalin et al., 1997).

Because the PMR includes both fetal deaths and deaths in the first week of life, it avoids conflicting judgments as to whether a fetus exhibited signs of life and variations in administrative practice regarding whether or not a death should be counted. In many countries, however, vital registration data are not sufficiently complete to allow reliable estimation of the PMR. Techniques now exist for collecting data on stillbirths, live births, and early neonatal deaths in population-based surveys (pregnancy histories) and applied in surveys including the DHS. However, there has been relatively less experience with pregnancy histories than with birth histories because of concerns about the quality of retrospectively reported pregnancy histories. Common problems with data quality include:

- Omission of stillbirths and early neonatal deaths;
- Difficulty in obtaining accurate information on gestational age or birth weight leading to the misclassification of stillbirths as late spontaneous abortions; and
- Heaping of the reported age at death of live births on 7 days, leading to the misclassification of early neonatal deaths as late neonatal deaths.[4]

---

[4] Heaping occurs when respondents do not know the exact age at death. Estimated ages at death are often reported on certain preferred ages, such as 7, 28, or 30 days, leading to a distorted age distribution of deaths in which too many deaths are reported at these preferred ages, and too few at the ages just before and after.

Prospective population-based surveys of pregnant women provide better quality data, but are expensive to undertake.

Evaluators typically calculate the PMR obtained from large population-based surveys at a national level and may aggregate data across countries to obtain a global or UN subregion statistic. Evaluators may also obtain sub-national estimates if sample sizes are sufficiently large.

The early neonatal component of the PMR may respond relatively quickly to programmatic interventions, for example, following the introduction of elements of the WHO "Essential Newborn Care Package." The stillbirth component may decline more slowly because it depends more on interventions that influence primarily maternal health and on the availability of technologies such as cesarian section. Survey-based estimates are generally subject to relatively large sampling errors, so detecting changes over short periods of time is impossible unless the changes are quite large. Also, retrospective survey-based estimates are often based on a five-year period prior to the survey. Therefore, evaluators should collect survey-based estimates of the PMR not more than every three to five years.

The following caveats bear mention. The PMR is sensitive to changes in the quality of data. For example, a rise in the PMR may indicate deterioration in perinatal outcomes, or it may indicate an improvement in the reporting of perinatal deaths. Therefore, an assessment of data quality is an essential component of analysis. In this context, evaluators often find it useful to separate the PMR into its two components: stillbirths and early neonatal mortality. Data quality is generally more problematic for stillbirths than for early neonatal deaths, because the problems of obtaining gestational age and ambiguity over the definition of stillbirths and fetal deaths are much less likely to be reported than deaths of live births (WHO, 1996b).

Evaluators should interpret facility-based estimates of the PMR with caution. The PMR in a facility is very sensitive to the types of deliveries occurring in the facility. Consequently, it may rise or fall in response to changes in the complexity of deliveries in the facility. In small facilities, the PMR will be very unstable because of the small number of deliveries and perinatal deaths; thus, the PMR is ineffective for monitoring change over time within the facility.

BIRTH WEIGHT SPECIFIC MORTALITY RATE (BWSMR)

Definition

The Birth Weight Specific Mortality Rate (BWSMR) is a stratification of a newborn mortality rate by birth weight grouping. (See indicator **Neonatal Mortality Rate – NMR.**) For example, the Birth Weight Specific Neonatal Mortality Rate for births over 2,500g is calculated as:

$$\frac{\text{\# of neonatal deaths weighing over 2,500 g at birth}}{\text{Total \# of live births weighing over 2,500 g at birth}} \times 100$$

And for births under 2,500g is calculated as:

$$\frac{\text{\# of neonatal deaths weighing under 2,500g at birth}}{\text{Total \# of live births weighing under 2,500g at birth}} \times 100$$

Evaluators can calculate birth weight specific mortality rates for perinatal deaths and stillbirths on the same basis.

Data Requirements

Number of deaths in a particular birth weight grouping and total number of births in the same weight grouping

Data Source(s)

Service statistics

HIS may collect data for this indicator in highly developed systems.

Purpose and Issues

As discussed in the preceding sections, birth weight is one of the most sensitive predictors of infant survival and is also a good predictor of maternal health and well-being. The mortality rate for low birth weight babies is much higher than for those with a normal birth weight. Stratifying newborn deaths by birth weight helps to determine the cause of death and therefore to identify where interventions are needed. For example, deaths of very small babies are more likely related to maternal causes predisposing to intrauterine growth retardation and preterm birth, whereas deaths of normal birth weight babies are more likely to be related to intrapartum asphyxia and poor obstetric care. In the first case, interventions should focus on the mother (improving nutrition and reducing antenatal infection) and, in the second case, should focus on improving the quality of delivery care. Evaluators can obtain additional information by stratifying birth weight by time of death (see Table III.E.4).

Evaluators can collect this type of indicator only in settings where all babies are weighed. It is therefore most appropriate for use in health facilities but has served in some community settings as part of a maternal and perinatal health care surveillance system (McCarthy, Lawn, and Ross, 2001).

One useful application of this type of disaggregation is to examine the number of intrapartum deaths in normal birth weight babies. If the quality of obstetric care is good (and women are not presenting very late in labor), then very few intrapartum deaths should occur because deliveries are expedited rapidly. The proportion of stillbirths in babies of normal birth weight may serve as a proxy indicator for intrapartum asphyxia and quality of delivery care.

Table III.E.4 Potential Causes of Death for Specific Age and Birth Weight Categories

Weight	Fetal Death	Intrapartum Death	Early Neonatal Death	Late Neonatal Death
Less than 2500g	Maternal infection, e.g. syphilis, other STIs Medical complications APH Hypertensive disease	Complications of preterm labor/IUGR Asphyxia	Complications of preterm labor/IUGR Infections	Infection, ARI Late complications of prematurity Tetanus
2500g and Above	Maternal infection, e.g. syphilis, other STIs, malaria Medical complications APH Hypertensive disease	Asphyxia and birth trauma Maternal infection	Asphyxia and birth trauma Infection	Infections, ARI Tetanus

This chart is a simplified representation of the BABIES (**B**irth weight, **A**ge at death, **B**oxes, **I**ntervention, **E**valuation **S**ystem). A more detailed explanation of this matrix and technique for interpreting the results appears in *The Healthy Newborn: A Reference Manual for Program Managers* (McCarthy, Lawn, and Ross, 2001).

- Percent of pregnant women who gain at least one kg per month in the last two trimesters of pregnancy
- Percent of non-pregnant women of reproductive age who have low body mass index (BMI)
- Percent of women with low mid upper arm circumference (MUAC)
- Percent of service delivery points with adequate supplies of mineral/vitamin supplements
- Percent of pregnant women who receive the recommended number of iron/folate supplements during pregnancy
- Percent of women of reproductive age with anemia
- Percent of women living in households using adequately iodized salt
- Percent of women who receive vitamin A supplementation in postpartum visits
- Percent of women with low serum vitamin A concentration
- Percent of women with night blindness in last pregnancy

Nutrition deficiencies diminish not only the individual woman's quality of life but also that of her children, family, and community because women are often income earners, food producers, and family caretakers. Adequate nutrition is vital both to the health and reproductive outcomes of women and to the health, survival, and development of their children. However, women's nutrition programs lack the resources often available to other nutrition and public health programs, in part because undernutrition in women often does not manifest itself through conspicuous outward signs. Nutrition often receives the lowest priority in terms of money spent (on either programming or evaluation), especially where nutrition interventions are embedded in a broader antenatal or MCH program. Often policymakers, program designers, and service providers direct their interest and efforts to "more pressing" issues.

Nutritional intervention programs tend to target the following three problems: (1) general nutritional deficiency (e.g., inadequate dietary intake), (2) specific micronutrient deficiencies, and/or (3) diseases directly affecting nutritional outcome (e.g., malaria, helminths). Intervention strategies addressing the first two problems include provision of supplements (food supplements, micronutrients), food production strategies, food-based strategies (genetic engineering, agricultural interventions), and dietary behavior change. Interventions to combat malaria and parasitic diseases include presumptive and therapeutic treatment. The indicators in this section reflect the tendency of nutritional interventions to date to promote women's nutrition as a means of improving pregnancy outcomes, rather than as an end unto itself. We anticipate certain adaptations in the indicators to monitor women's nutrition programs, as interventions increasingly focus on the nutritional status of women – because of the benefits improved nutrition will have on the well-being of the individual woman as well as of her child(ren).

At the program level, there is growing awareness that health and nutrition programs implemented well before women become pregnant will have long-term impacts on both the mother and child, one reason for which international donor agencies have demonstrated a renewed interest in women's nutrition and nutrition education. Intervention strategies need to go beyond the conventional approach of providing services to pregnant women and mothers through traditional maternal and child health care programs; they must also take advantage of the opportunities presented through community-based approaches (village volunteers, community events as a venue for health and nutrition messages) and must link nutrition with initiatives implemented in other sectors (e.g., agriculture extension, education, micro-credit).

This *Compendium* presents women's nutrition as a discrete topic. Yet at the field level, programmatic activities involving women's nutrition are usually integrated ("bundled") with other health services. Indeed, women's nutrition is intricately linked to other health outcomes. For example, when parasitic diseases and malaria diminish the nutritional status of women, the health status of the newborn suffers. The natural link between women's nutrition and safe motherhood is further enforced by the programmatic reality that prenatal visits allow the public health establishment to reach pregnant women of reproductive age with other health interventions. A woman's nutritional status also influences the body's ability to fend off infections, including opportunistic infections associated with HIV. The prevalence of AIDS in a given area in turn affects women's nutritional status, if the population becomes so depleted that it cannot accomplish basic household activities because of illness, fatigue, loss of income, and increased medical expenses.

Nutrition also plays a key role in adolescent health programs. Nutritional intake during childhood and adolescence directly affects health and well-being at all stages in life. Well-nourished girls perform better in school and have a greater capacity for physical activity than do undernourished girls. Adequate nutrition in childhood also determines whether a young woman will reach her own childbearing years with adequate weight and nutritional status to produce a healthy child. Moreover,

nutrition-related interventions appear to be more effective in changing behavior among younger than older women; thus, investments in youth programming may lead to more sustainable behavior change over time.

This section focuses on two types of results: outputs measured at the program level (e.g., results among clients or other facility-based data) and outcomes measured at the population level. Because the latter requires data from Multi-Indicator Cluster Surveys (MICS), Knowledge, Practices, and Coverage (KPC) surveys, and DHS-type or other representative surveys of the population in the area of project activity, organizations with limited human or financial resources for program evaluation may find it unfeasible to gather population-based data. As is true in other areas of reproductive health, program-based and population-based data can provide very different information regarding coverage. For example, the different data may simultaneously reveal a high level of supplementation coverage of pregnant women attending prenatal clinics, but a very low coverage of pregnant women in the population (if relatively few pregnant women attended prenatal care). For the indicators mentioning both program- and population-based data as sources of data, this caveat warrants attention.

The section does not explicitly cover the indicators most useful for diagnostic or screening purposes (i.e., to identify vulnerable populations in need of food supplementation), although two indicators – **Percent of Non-pregnant Women of Reproductive Age Who Have Low Body Mass Index (BMI)** and the **Percent of Women with Low Mid Upper Arm Circumference (MUAC)** – can serve this end. We include two biological markers – serum vitamin A and hemoglobin – that programs with budgets sufficient to collect such data may find useful to measure. However, many programs may opt to omit them, because they can be more expensive, time-consuming, and logistically difficult to collect than other nutrition indicators. Different intervention priorities, such as the use of multi-vitamins or calcium supplements, are expected to gain increasing prominence in the coming years. However, because these interventions remain largely experimental, we have omitted indicators for them in the *Compendium*.

Women's nutrition differs programmatically from other RH areas in that the causal chains are longer and more complex, in part because of confounding biological factors. As such, reaching members of the intended audience with the intervention does not guarantee the expected effect. In family planning programs, one assumes that when a woman uses an effective method of contraception correctly, she will avoid pregnancy. By contrast, in the case of women's nutrition, a thin woman (and her thin child) may need food energy but fail to respond to increased food intake because of HIV infection, TB, or malabsorption. Thus, when program planners design and evaluators assess programs, they must specify intended effects and the size of the expected effects, taking into consideration contextual factors. The selection of indicators then follows.

The selection of indicators depends upon the type of program: national versus small area, formative/pilot versus ongoing. In the area of women's nutrition, formative/pilot studies continue to play an essential role in understanding the multi-causal nature of most nutrition problems, and indicators based on biochemical data can be very useful in this context. Such measures may be less practical for national programs. However, the DHS has demonstrated the feasibility of collecting samples using the finger-prick technique to measure the prevalence of anemia at the national level.

Where possible, evaluators should look for data that already exist or will be collected as part of larger surveys using widely accepted data-collection methods, such as the DHS, MICS, and KPC surveys. While very little research has tested the validity and responsiveness of women's nutrition indicators, a recent analysis of data from the Philippines (Adair, 1998) demonstrates the utility and importance of this type of research.

The inclusion of women's nutrition in the *Compendium* reflects a growing interest in and recognition of the importance of this subject among reproductive health professionals. As programs designed to improve women's nutrition become more widespread, the science for evaluating them will advance accordingly. In anticipation of greater interest in evaluation in this area, we outline a number of the methodological challenges for evaluating women's nutrition programs.

Methodological Challenges of Evaluating Women's Nutrition Programs

- **Nutrition is a complex area to evaluate because of the multiple determinants.**

Numerous other sections of the *Compendium* discuss the problems of establishing a causal relationship between the intervention in question and the desired outcome. This challenge applies equally to programs on women's nutrition because of the multi-sectoral nature of determinants of poor nutrition as well as because of the interventions necessary to improve nutritional status. That nutritional interventions are rarely carried out in isolation of other health programs, further complicates establishing a definitive cause-and-effect relationship.

- **The cut-off points for indicators of a woman's nutritional status can vary according to the period within the reproductive cycle (i.e., during pregnancy and lactation).**

The ideal indicator of women's nutritional status would have the same cut-off points or, at least, cut-off points clearly identified for different periods during the reproductive cycle for all women of reproductive age in a given population. However, this is often not the case. For example, there is no clear agreement on the cut-off points for body mass index (BMI) during pregnancy and lactation. Furthermore, applying the correct cut-off points depends on determining whether a woman is pregnant or not; such a determination is sometimes impossible in a program setting.

- **Because programs with explicit objectives to improve women's nutritional status have been rare, program evaluation in this area is also relatively underdeveloped.**

The field of program evaluation in family planning developed in large part in response to critics who cited the millions of dollars being spent in this area and who questioned the effectiveness of programs to achieve results. In contrast, in terms of women's nutrition, a fair amount of work has demonstrated the efficacy of specific treatments, but far less evaluation has been conducted on the effectiveness of nutritional interventions at the program level. One possible reason is that women's nutrition programs have not been challenged to demonstrate their effectiveness in terms of health outcomes (although the evaluation of micronutrient interventions is somewhat an exception to this generalization). Moreover, given the relatively low levels of funding in this area, few have advocated using limited resources on sophisticated methods of evaluation. The trend has been to invest available funds in program activity rather than data collection and analysis.

Much of the evaluation of women's nutrition interventions to date has consisted of process and output type measures (e.g., number of talks given, number of women visiting the center, number of food supplements distributed). Where resources are limited, these measures are useful to track program activity, and (in the case of micronutrient supplements) they serve as proxies for population-level measures especially in the absence of outcome measures. However, these measures of process and output are only as good as the management information systems that generate them. The indicators presented in this *Compendium* move beyond process to the results expected from these programs, either at the program or population level.

Organization of this Section

As mentioned above, women's nutrition programs aim to (1) improve overall nutrition status by increasing caloric intake, decreasing energy expenditure, or ensuring better care, (2) provide micronutrient supplements or access to fortified foods – primarily iron, iodine, and vitamin A (and more recently multiple micronutrient supplements), and (3) manage diseases directly affecting nutritional outcomes, such as malaria and helminths. The beneficiaries of these efforts tend to be eligible women (and their children) in a defined geographical or administrative area or to be clients of a particular NGO. The indicators in this section provide useful information on:

1. The performance of the program in delivering these services (outputs measured at the program level);

2. The results achieved in terms of nutritional status
    (a) in women participating in the program (outputs measured at the program level), and/or
    (b) among relevant subgroups of the general public (outcomes measured at the population level, through representative sample surveys).

The **performance** indicators in this section reflect common interventions for women's nutrition programs focusing on women during pregnancy and the post-partum period.

Interventions for pregnant women include:
- Iron/folate supplementation (and where applicable, multiple micronutrient supplementation);
- Malaria prophylaxis;
- Anthelminthic treatment; and
- Increase in food intake.

Interventions for women in the post-partum period include:
- Vitamin A supplementation;
- Iron/folate supplementation; and
- Increase in food intake.

Indicators to measure the **results** of these interventions among clients and/or relevant subgroups of women in the general population include:

All women:
- Body mass index (to detect thin and overweight women);
- Mid-upper arm circumference (MUAC);
- Availability of iodized salt in the household; and
- Serum retinol.

Pregnant women:
- Weight gain of at least one kilo/month in last two trimesters of pregnancy; and
- Night blindness (indicative of vitamin A deficiency).

Table III.F.1 provides an overview of this section. The indicators are ordered by type of intervention (programs to increase caloric intake and micronutrient supplementation including iron, iodine, vitamin A as well as iodine fortification). The indicators on prophylaxis or presumptive treatment for malaria and helminths are relevant in this section, but appear under Part III.D on Safe Motherhood. Within each type of intervention, those that measure program performance are presented first, followed by those that measure nutritional status among clients and/or subgroups of women in the general population.

Table III.F.1 Women's Nutrition: Types of Interventions, Indicators of Program Performance, and Indicators of Nutritional Status

Type of Intervention	Measures of Program Performance (titles abbreviated)	Results in Terms of Nutrition Status (titles abbreviated)
Interventions to increase energy intake		Weight gain of at least one kg/month in pregnant women Body Mass Index (overweight, underweight) Mid upper arm circumference
Micronutrient Supplementation	Adequacy of supplies at SDP (iron, iodine, Vitamin A)	
• Iron	Iron folate supplementation during (last) pregnancy	Percent of women with anemia
• Iodine	Percent of women living in households with iodized salt	Percent of women with low-serum Vitamin A concentration
• Vitamin A	Vitamin A supplementation in postpartum visits	Percent of women with night blindness in last pregnancy
Management of diseases directly affecting nutritional outcomes		
• Malaria	Percent of women receiving treatment during pregnancy	
• Helminths	Percent of women receiving treatment during pregnancy	

PERCENT OF PREGNANT WOMEN WHO GAIN AT LEAST ONE KG PER MONTH IN THE LAST TWO TRIMESTERS OF PREGNANCY

Definition
The percent of women gaining at least 1.0 kg per month in the second or early third trimester of pregnancy (Krasovec and Anderson, 1991a and 1991b)

This indicator is calculated as:

$$\frac{\text{\# of women gaining at least 1.0 kg per month in the second and third trimesters of pregnancy}}{\text{Total \# of pregnant women}} \times 100$$

Data Requirements
Two or more recordings of weight after the third month of pregnancy

Data Source(s)
Service statistics, prenatal cards, or other clinic-based records; sample of home-based records reviewed

Purpose and Issues
This indicator measures weight gain during pregnancy, one of the most critical factors in determining both birth outcomes and maternal nutritional outcomes of pregnancy. Weight gain is particularly important for women who are underweight prior to pregnancy and for women who are pregnant during times of acute nutritional stress, such as famines or seasons of food scarcity. Underweight women (Body Mass Index <18.5) need to gain between 12.5 and 18 kg during pregnancy in order to lower their risk of producing low-birth-weight (LBW) babies (IOM/NAS, 1990). Average weight gains for women in developing countries (5-9 kg) are much lower than these recommendations, and much lower than averages for developed-country women (10.5-13.5 kg). At a minimum, women should gain at least 1.0 kg/month during the last two trimesters. A WHO report (1995a) indicates that a higher gain of 1.5-2.0 kg per month improves infant outcomes (LBW and IUGR).

Low weight gain during pregnancy is associated with LBW, intrauterine growth retardation (IUGR), gestational duration, fetal and neonatal mortality, and maternal nutritional status postpartum.

This indicator's strength is that it reflects the importance of routine and high-quality antenatal care through multiple prenatal visits. Moreover, it focuses the attention and care of both the health worker and the woman on weight gain and weight gain promotion rather than simply on determining maternal nutritional status at any one point in time.

This indicator's major limitation is that the population covered may not fully represent the intended population, because a very small percentage of women in many developing countries routinely attend prenatal services. Frequent attendees tend to be either women with pregnancy complications or women of higher socio-economic and educational status. The difficulty in monitoring maternal nutrition during pregnancy is that many women do not get antenatal care, or they have only one visit late in the pregnancy.

This indicator is most often used by NGOs or PVOs working in a limited geographic area. Evaluators face difficulty obtaining this information from large public health centers in developing countries that keep antenatal records. An adult scale and antenatal cards are essential to obtain this information.

PERCENT OF NON-PREGNANT WOMEN OF REPRODUCTIVE AGE WHO HAVE A LOW BODY MASS INDEX (BMI)

Definition
Low Body Mass Index (BMI), the ratio of weight to height (kg/m^2), measures chronic energy deficiency or "thinness" in non-pregnant women

The standard cut-off for non-pregnant, non-lactating women aged 15-49, determined by the International Dietary Energy Consultative Group, is a BMI of 18.5. Further refinements in levels of CED are:
- Grade I: 17-18.4 (mild);
- Grade II: 16-16.9 (moderate); and
- Grade III: <16 (severe)

(James et al., 1988).

This indicator is calculated as:

$$\frac{\text{\# of non-pregnant women with a BMI below 18.5}}{\text{Total \# of non-pregnant women between the ages of 15-49}} \times 100$$

Data Requirements
Weights (in kilos) and heights (in meters squared) of non-pregnant women of reproductive age

Data Source(s)
Population-based surveys

Purpose and Issues
The advantage of BMI, a well-accepted measure of energy deficiency among women, over weight for height as a measure of thinness is that it does not require reference tables for interpretation. However, it may present difficulties to some field workers in service delivery programs because of the mathematical calculations required. Tools (e.g., tables, wheels) have been developed to assist with these calculations.

Rapid changes in anthropometric measures as a result of the adolescent growth spurt complicate assessing the nutritional status of those below 18 years of age (i.e., it increases the variance in BMI). Despite this caveat, BMI is nonetheless recommended for use with adolescents. A related (additional) indicator to BMI is the woman's weight, which reflects both acute and chronic nutritional stresses. The cut point for identifying women who are undernourished is 45 kg (ACC/SCN, 1992).

Because BMI varies with body shape or the Cormic index (sitting height divided by standing height), some have argued that data on sitting height should be collected where possible and that the BMI should be adjusted for the Cormic index. However, others consider this adjustment to be impractical, given that the calculation of BMI itself is methodologically challenging to some field workers.

Evaluators should use caution when they interpret adult anthropometric results because of the lack of validated outcome data for interpreting the results. The evaluators should disaggregate data by age and lactational status.

Although this indicator specifies non-pregnant women, BMI also is commonly used to identify women who need to gain more weight during pregnancy in order to improve infant outcomes of pregnancy (low birth weight, intrauterine growth retardation, and perinatal mortality). It is also used to monitor women during pregnancy.

An alternative indicator to BMI in situations where it is impractical to get weight and height data is that for middle-upper arm circumference (MUAC), which is based on a single anthropometric measure. (See next indicator, **Percent of Women with Low Mid Upper Arm Circumference**).

BMI is also useful for identifying women who are overweight (BMI > 25.00 kg); a BMI > 29.00 is considered obese (IOM, 1990). The prevalence of overweight women is increasing rapidly in certain developing countries. However, we are not aware of large-scale interventions to reduce the percentage of overweight women in a developing country which have been evaluated

based on BMI. For this reason, we do not include BMI to measure overweight as a separate indicator in this *Compendium*.

Gender Implications of this Indicator

Limiting food intake during pregnancy is a gender-based, harmful cultural practice theoretically linked to the idea that limiting weight gain will limit the infant's head circumference, so that the birth will be less difficult. The practice primarily occurs in settings such as South Asia, where women eat last and least, even when not pregnant. Thus, women frequently enter pregnancy undernourished and become more so throughout pregnancy. In reality, undernourished pregnant women are at much greater risk of poor birth outcomes than are nourished women. They are more likely to be vitamin A-deficient and to be anemic, both of which also increase the risk of maternal and fetal morbidity and mortality. Some nutritionists believe it beneficial to limit weight gain during pregnancy; also, they may use this rationale to save face for being poor and unable to eat more. More efforts are needed to educate husbands, mothers-in-law, and communities that pregnant women must eat more, not less, and that nutritious foods benefit both mother and fetus and lead to better birth outcomes.

PERCENT OF WOMEN WITH A LOW MID-UPPER ARM CIRCUMFERENCE

Definition

The percent of women with a middle upper arm circumference (MUAC) below 22.5 cm (ACC/SCN, 1992)

This indicator is calculated as:

$$\frac{\text{\# of women with a mid upper arm circumference below 22.5 cm}}{\begin{array}{l}\text{Total \# of women between the}\\\text{ages of 15-49}\end{array}} \times 100$$

Data Requirements

A measure of MUAC in women of reproductive age (15-49)

Data Source(s)

DHS or other population-based surveys; the KPC_{2000} survey including the collection of data on the percentage of mothers with children under the age of two years who have a low MUAC

Purpose and Issues

MUAC is an anthropometic measure used primarily for screening, because it changes slowly in large populations. However, it is potentially useful for evaluating the impact of interventions in a given (limited) population.

The measure of mid upper arm circumference is a useful anthropometric measure because it's easily obtained in clinical settings or during population-based surveys. The tapes are portable and inexpensive, and persons with limited education (e.g., community workers, TBAs) can learn to take this measurement accurately. The measurement of MUAC not only yields useful data but also raises awareness about nutritional status among those participating in the study. In settings with limited infrastructure and resources, MUAC may be the only feasible anthropometric indicator to use. An additional advantage is that the same cut-off value can be used to define undernutrition in both pregnant and non-pregnant women because values change only slightly during pregnancy.

MUAC is primarily used for screening rather than evaluation purposes. MUAC is correlated with pre-pregnancy weight and may be useful for identifying pregnant women at risk of IUGR, especially where scales are not available (WHO, 1995a). Where used, the data should be disaggregated by age and reproductive status.

Cut-off values between 21.0 and 23.5 cm are consistently related to biological risk of LBW and fetal and infant mortality in Asia and Latin America; data are not available for Africa. Further validation of MUAC is needed.

PERCENT OF SERVICE DELIVERY POINTS WITH ADEQUATE SUPPLIES OF MINERAL/VITAMIN SUPPLEMENTS

Definition
"Adequate supply" is the availability and quality of mineral/vitamin supplements (iron, iodine, and vitamin A) at the service delivery point (SDP) at the time of data collection. To compute adequacy, evaluators determine the number of individual doses (daily or otherwise) of acceptable quality supplements relative to the client population served. The evaluators should calculate each type of supplement (iron, iodine, and vitamin A) separately, because the frequency of doses and, therefore, the amount necessary depend on the type of supplement (i.e., daily iron supplements vs. single postpartum dose of vitamin A).

The quality of the mineral/vitamin supplement supply (iron, iodine, folate, and vitamin A) is acceptable if the supplements are labeled properly, not expired, and are stored under the recommended climatic and lighting conditions. The evaluators should assess adequacy of each type of supplement separately, because some products are often difficult to procure in a given country (e.g., iron/folate) in contrast to others more readily available (e.g., Vitamin A, provided by UNICEF).

This indicator is calculated as:

$$\frac{\text{\# of SDPs with an adequate supply of quality mineral/vitamin supplements}}{\text{Total \# of SDPs}} \times 100$$

Data Requirements
A count of the number of SDPs in the catchment area; a count of the potential client population(s) in the catchment area served at each SDP; a count of units of each supplement listed by form of the supplement (e.g., iron: tablets and drops; iodine: tablets and injectables; vitamin A: high- and low-dose capsules) of acceptable quality at the SDP; the volume of supply of each mineral/vitamin in terms of "individual doses;" and number of doses of each supplement judged (a) to be sufficiently well stocked and (b) of adequate quality (see operational definitions below).

Data Source(s)
Program records indicating the number of SDPs and the population in the catchment area; and inventory of each SDP (special study) and inspection of each unit of supplement to determine the number of supplements that are of acceptable quality

Purpose and Issues
This indicator is important at the program level to evaluate the extent SDPs have supplements that are both available and of acceptable quality to meet clients' nutritional needs. For this indicator to be useful, evaluators must define the measurement of "adequate" and "of sufficient quality." One measures sufficient quantity by estimating the size of the catchment area and the subgroup within that area in potential need of the supplement. One then calculates the average quantity of each supplement needed per recipient in the target population for a specific reference period. This approach allows a crude calculation of the "sufficient quantity" for each supplement. There is no universally accepted standard for measuring adequate supplies; however, evaluators should consider the type of supplement, the frequency of supply, and the amount of supply available at the SDP when they define adequacy.

Certain criteria for quality apply to all three supplements:
- Supplements should be properly labeled (supplement name, volume, usage, dosage, medical contraindications, and expiration date);
- Supplements should not have not expired; and
- Supplements should be stored in a cool, dry place, under conditions specified by the manufacturers.

Additional criteria for quality also apply to specific supplements, as follows:

- Iron: Iron tablets and drops are considered acceptable if at least 90 percent of the tablets in the bottle are intact, and if any other recommendations from the manufacturer on proper

storage are being followed; and

- Vitamin A: The quality of vitamin A supplements is considered acceptable if supplements are stored away from light between 0°C and 30°C, and liquid vitamin A is discarded if it has been open for more than two months.

This indicator requires that the supplements meet both criteria: of adequate quantity and of sufficient quality. Thus, evaluators assess the results of these two factors simultaneously to determine if a given SDP has an "adequate supply."

This indicator measures the presence of products at service delivery facilities. It does not, however, measure the effective distribution of these products to the intended beneficiaries. Staff awareness, motivation, and training will strongly influence this process.

An additional process indicator of the adequacy of supply is the frequency of stockouts (i.e., the percentage of SDPs that experience a stockout of supplements at least once over a 12-month period).

For an additional discussion on indicators for Commodities and Logistics, see Part II.E.

PERCENT OF PREGNANT WOMEN WHO RECEIVE THE RECOMMENDED NUMBER OF IRON/FOLATE SUPPLEMENTS DURING PREGNANCY

Definition
The percent of women who receive iron/folate supplements in accordance with local policy or protocols

USAID (1999) recommends at least 90 tablets of iron/folate during pregnancy. Alternatively, the IVACG/WHO/UNICEF (1998) recommendations indicate that pregnant women should receive iron/folate supplements for at least six months of pregnancy (and for an additional three months postpartum if pregnancy anemia prevalence is > 40 percent).

This indicator is calculated as:

$$\frac{\text{\# of pregnant women who receive iron/folate tablet}}{\text{Total \# of pregnant women}} \times 100$$

Data Requirements
Information on the number of pregnant women who were issued iron/folate tablets during last pregnancy; the number of tablets issued; and the total number of women who gave birth in the reference period

Data Source(s)
Program statistics (most common source) or population-based surveys

Purpose and Issues
This indicator measures whether women who give birth in a given reference period receive the minimum number of iron/folate supplements in the form of tablets, based on local policy or international standards. If the source of data is a population-based survey, the evaluator should calculate the indicator for the last pregnancy.

Worldwide, iron deficiency is the most common nutrient deficiency, and pregnant women are especially vulnerable. Pregnant women need iron to support their enlarged blood volume, to provide for placental and fetal needs, and to replace blood loss in childbirth. The fetus relies on maternal iron stores to create adequate reserves of its own, which in tandem with the highly recessible iron in breast milk, will meet the iron needs of the normal birth weight infant through the first six months of life. Iron supplementation is particularly recommended during the second and third trimesters when iron stores become depleted over the course of pregnancy (Whitney, Cataldo, and Rolfes, 1998).

Anemia is defined as abnormally low hemoglobin concentration. The results are fatigue, weakness, headaches, apathy, pallor, and poor resistance to cold temperatures. Since pregnant women are especially vulnerable, providing iron/folate supplements is particularly important. Supplementation early in the pregnancy is desirable, particularly where deficiency levels are high.

Research also suggests that folate supplements taken at least one month prior to conception and continued throughout the first trimester of pregnancy can prevent neural tube defects. Such defects cause serious disabilities and infant mortality and commonly arise in the first weeks of pregnancy before a woman may realize she is pregnant.

This indicator captures the distribution aspect of iron/folate supplements, but not the actual consumption. Clients must receive appropriate counseling on why and how to take iron/folate supplements.

An alternative indicator that reflects the adequacy of the program in meeting the needs of specific clients is

- Number of tablets distributed per eligible client.

PERCENT OF WOMEN OF REPRODUCTIVE AGE WITH ANEMIA

Definition

The percent of women of reproductive age who have anemia, which is an inadequate level of hemoglobin

WHO (2000b) has defined anemia as mild, moderate, or severe based on the following cutoff values (g/dl) for hemoglobin level:

	Mild	Moderate	Severe
Pregnant	10-10.9	7.0-9.9	<7.0
Non-Pregnant	11-11.9	8.0-10.9	<8.0

In short, pregnant women with a hemoglobin level less than 11g/dl and non-pregnant women with a level less than 12g/dl are considered anemic.

Note: Evaluators may use mean hemoglobin (a continuous variable) instead of the above categories of mild, moderate, or severe (the same information in categorical form).

This indicator is calculated as:

$$\frac{\text{\# of women who have anemia}}{\text{Total \# of women between the ages of 15-49}} \times 100$$

Data Requirements

Hemoglobin concentration measures on a sample of women of reproductive age (or of women included in a surveillance system), including both pregnant and non-pregnant women

Data Source(s)

Population-based surveys or surveillance

Purpose and Issues

Anemia is a condition in which an inadequate number of red blood cells or an inadequate amount of hemoglobin prevents the body from functioning properly. Hemoglobin is a protein in red blood cells that carries oxygen to the brain, muscular system, immune system, and other parts of the body. Inadequate oxygen reduces the physical and mental capacity of individuals.

Over 40 percent of non-pregnant and 56 percent of pregnant women in lesser developing countries are anemic (ACC/SCN, 2000b). In industrialized countries anemia also affects women, especially those of lower socio-economic status. Iron deficiency is the primary cause of most anemia in poor environments. Progressive loss of iron stores leading to functional-tissue iron deficiency precede the occurrence of iron deficiency anemia. In addition to iron, other nutritional deficiencies (e.g., folate, vitamin B-12, and vitamin A) can cause anemia, as can non-nutritional factors such as acute and chronic infections (malaria, hookworm, HIV) and genetic conditions such as thalassemia and sickle cell trait.

Various factors may influence estimates of anemia prevalence, including sex, age, pregnancy status, and altitude; thus, evaluators must adjust individual level data for these factors. Among women of reproductive age, adolescent girls and pregnant women are at most risk for anemia: adolescents because of the onset of menstruation and pregnant women because of the increased blood volume associated with pregnancy. Severe anemia among pregnant women resulting from iron deficiency is associated with an increased risk of maternal and fetal mortality and morbidity and of intrauterine growth retardation (WHO, 2000b).

Additional laboratory tests, such as measurement of serum ferritin and/or malarial and parasitic egg counts, are necessary to determine if iron deficiency is the primary "cause" of the anemia. However, these tests are frequently impractical for field-based use. Until a simple, cost-effective test for measurement of iron deficiency is widely available and feasible for program application, the prevalence and distribution of anemia will continue to be used to estimate the extent, trends, and severity of both anemia and iron deficiency anemia at the population level.

This indicator is also useful for monitoring sub-populations (e.g., pregnant women, lactating women, women who receive antenatal care, women who receive postpartum care) and for evaluating interventions directed

towards these subgroups. Data should be disaggregated
by age and reproductive status.

Gender Implications of this Indicator

There may be gender-related food taboos that con-
tribute to high levels of anemia by denying women
iron-rich foods. In addition, social norms may dic-
tate the order in which family members eat, thus lim-
iting women's access to iron-rich foods. It may be
difficult for women to obtain iron folate supplements
if they lack freedom of movement to travel to distri-
bution points or lack access to household financial
resources for transportation to distribution points or
to purchase commodities.

PERCENT OF WOMEN LIVING IN HOUSEHOLDS USING ADEQUATELY IODIZED SALT

Definition

This indicator, measuring the percent of women who live in households with iodized salt, is a proxy measure for the number of women who may be receiving adequate amounts of iodine.

"Adequately iodized salt" is defined as salt containing 15+ ppm of iodine.

This indicator is calculated as:

$$\frac{\text{\# of women who live in households with salt containing 15+ ppm of iodine}}{\text{Total \# of women}} \times 100$$

Data Requirements

Results of testing household salt used for cooking and/or as table salt

Data Source(s)

Population-based household surveys; the testing of households using iodized salt is part of the core questionnaire in the DHS and in other surveys such as MICS.

Purpose and Issues

Iodine deficiency disorders (IDD) are prevalent throughout the world and affect over 570 million people, mainly in developing countries (WHO, 1993). IDD interventions often focus on women of reproductive age because of their increased need for iodine during pregnancy. Iodine deficiency in pregnancy may impair the development of the fetus, and thus may cause extreme and irreversible mental and physical retardation known as cretinism (Whitney, Cataldo, and Ross, 1998). Mild deficiency is very common and probably has harmful effects.

The purpose of this indicator is to evaluate the availability of adequately iodized salt in a given population. Salt is the main food seasoning among people all over the world. Iodization of salt, therefore, is the most effective means of ensuring that a population receives adequate amounts of iodine. Moreover, iodizing salt is relatively easy and low cost.

This indicator is a proxy for iodine status, given the difficulty of obtaining the latter from large-scale surveys. Where the data permit, disaggregating by geographical/ecological zone and by socio-economic level is useful.

PERCENT OF WOMEN WHO RECEIVE VITAMIN A SUPPLEMENTATION IN POSTPARTUM VISITS

Definition

The percent of breastfeeding and non-breastfeeding women who receive two high dose supplements (200,000 IU per dose) of vitamin A within six weeks of giving birth

This indicator is calculated as:

$$\frac{\text{\# of women receiving two high-dose supplements within six weeks of delivering}}{\text{Total \# of women who deliver within a given reference period}} \times 100$$

Data Requirements

The total number of births during a given reference period and the number of women receiving two high-dose vitamin A supplements within six weeks of delivering

Data Source(s)

Program statistics (usual source) or population-based surveys, such as MICS, DHS, and KPC (possible alternative source)

Purpose and Issues

Vitamin A supplementation during lactation raises (and maintains) the concentration of vitamin A in the breast milk of women with vitamin A deficiency. Mega-doses of vitamin A, however, can potentially harm a fetus; therefore, women who could become pregnant must not receive vitamin A.

Different expert groups differ on the criteria for the "safe" infertile period after delivery during which a relatively high dose of Vitamin A supplement may be given. For example, the 1998 WHO/MI document on Safe Vitamin A dosage during pregnancy and lactation recommends that, in hyperendemic vitamin A-deficient areas, breastfeeding mothers receive 200,000 IU vitamin A within eight weeks of delivery – provided the woman is not pregnant. Non-breastfeeding women can be safely supplemented within six weeks of delivery. This level of supplementation will raise and maintain the vitamin

A content of breast milk and will offset the depleting effect lactation may have on the mother's own vitamin A stores (ACC/SCN, 1994). The IVACG Informal Consultation on Vitamin A Supplementation, Yverdon, Switzerland, recommends a higher dose (400,000 IU), preferably given in two doses within six (for non-breastfeeding mothers) to eight weeks (for breastfeeding mothers). To avoid confusion among health personnel about the safe infertile period, PAHO currently advises that all mothers take two doses of the supplement (200,000 IU per dose and at least 24 hours between doses) within six weeks postpartum (PAHO, 2001). This recommendation is also advised by UNICEF.

Evaluators usually calculate this indicator from service statistics, but they can obtain it for the general population from population-based surveys. Evaluators should disaggregate findings for lactating versus non-lactating women (to ensure that the program is reaching both groups), also by urban/rural or by socio-economic level, if the numbers permit.

One potential problem in the calculation of this indicator is the clients may deliver at a different place from the one where they receive the supplementation. If the indicator is based on an overall figure for a district, it is generally more accurate than if it were based on the data from specific clinics. Similarly, it is essential to specify whether this indicator measures supplements distributed through outreach workers to mothers delivering at home, or only those given at service delivery points.

Evaluators can adapt this indicator so that it refers to all women, not just to those in the post-partum period, to evaluate interventions aimed at all women through programs such as "Healthy Days" or "National Immunization Days."

An alternative indicator reflecting the adequacy of the program in meeting the needs of specific clients is the number of capsules distributed per eligible client.

PERCENT OF WOMEN WITH LOW SERUM VITAMIN A CONCENTRATION

Definition
The percent of women whose serum vitamin A (retinol) is less than 1.05 umol/l

This indicator is calculated as:

$$\frac{\text{\# of women with serum vitamin A } <1.05 \text{ umol/l}}{\text{Total \# of women}} \times 100$$

Data Requirements
Levels of retinol in serum (note: plasma levels give comparable results, WHO, 1996c)

Data Source(s)
Population-based surveys

Purpose and Issues
Serum retinol has been the indicator used most often in making a biochemical assessment of vitamin A status. Methods are now available to measure serum retinol with only a finger stick (HPLC on 50 uL of serum and dry-blood-spot techniques are in use; other methods are under development).

As for validity, the relationship between serum retinol and vitamin A status as indicated by total-body reserves is complex and non-linear. Vitamin A circulates in blood as retinol bound to its specific carrier protein, retinol-binding protein (RBP). The level of retinol in the blood is under homeostatic control over a broad range of body stores and reflects body stores only when they are very low or very high. Since RBP is an acute phase protein, acute and chronic infections can make interpretation of serum retinol levels difficult. Thus, serum concentration is not a valid indicator of vitamin A deficiency in individuals, but a frequency distribution of serum retinol concentrations can be informative for populations (WHO, 1996c).

Collecting blood samples is clearly essential for using this indicator. Because the ease of collecting blood samples varies by setting (e.g., it is particularly difficult in populations with high prevalence of HIV), the practicality of this indicator is limited.

Cut-off points for serum retinol to indicate vitamin A deficiency have been established more firmly for children than for women. The cutoff point used most commonly for adults is 1.05 umol/l, although some have used the same cut-off point recommended for children — 0.7 umol/l. The major justification for this cutoff point is based on data from the NHANES survey of the US population presented by Pilch (1987). This study suggests that serum retinols increase with age; for those aged 8-74 years, concentrations between 0.7 and 1.05 umol/l may improve with increased consumption of vitamin A, and some individuals with these concentrations may exhibit impairment of function. Evaluators may find it prudent, depending on the purpose of the survey, to present data using both cutoff points and so allow comparisons with the results of almost all surveys.

An alternative measure of vitamin A in lactating women is based on vitamin A concentration in breast milk. Breast milk retinol is very useful in evaluating vitamin A interventions because it has been shown to be the biochemical indicator most sensitive to measuring the impact of vitamin A interventions (Stoltzfus and Underwood, 1995). Logistical difficulties in maintaining the sample under the necessary temperature conditions make it less feasible for use in the context of large-scale population surveys. Thus, vitamin A concentration in breast milk is not included as a separate indicator in this *Compendium*.

PERCENT OF WOMEN WITH NIGHT BLINDNESS IN LAST PREGNANCY

Definition
The percent of women who had night blindness during the last pregnancy

Maternal night blindness is marked by impaired scotopic (adjusting to dim light) vision during pregnancy that commonly recurs during repeated pregnancies and occasionally extends into the postpartum period (Christian et al., 1998a).

A prevalence rate of night blindness in pregnant women higher than five percent signals that vitamin A deficiency is a problem of public health significance in that population.

This indicator is calculated as:

$$\frac{\text{\# of women who had night blindness during the last pregnancy}}{\text{Total \# of ever-pregnant women}} \times 100$$

Data Requirements
Self-report of the condition during last pregnancy. In areas where night blindness exists, qualitative research needs to be conducted to determine the local term or description of symptoms for night blindness in that area. Evaluators need to distinguish impaired vision in dim light from impaired vision in daylight.

Data Source(s)
Population based surveys (e.g., DHS, MICS)

Purpose and Issues
Maternal night blindness is an indicator of severe vitamin A deficiency. Information on the validity of the indicator comes mainly from a study of low-dose supplementation with vitamin A or B-carotene to women of reproductive age in Nepal (West et al., 1999).

Supporting its validity as an indicator of maternal vitamin A status, night blindness in pregnancy was strongly

associated with low retinol in serum and breast milk, abnormal conjunctival cytology, and impaired dark adaptation (Christian et al., 1998a). Furthermore, the incidence of maternal night blindness was reduced by two-thirds with vitamin A interventions, providing causal evidence that the night blindness resulted from vitamin A deficiency (Christian et al., 1998b).

Night blind women carry considerably greater risks for health and survival than non-night blind women do. In the Nepal study, pregnant women reporting night blindness were more likely to be anemic, ill, acutely undernourished and to be consuming a nutritionally poorer diet (Christian et al., 1998a). They were also at higher risk of mortality than others were (Christian et al., 2000). Maternal night blindness was associated with a four-fold higher risk of all-cause mortality for up to two years following the night blindness.

A history of night blindness is easy to obtain when a local term exists for the condition, but interviewers must ask questions concerning the history in a standardized format. Those analyzing the data must exclude cases of night blindness that report daytime vision problems. Because night blindness tends to occur in the later part of pregnancy, surveys that measure night blindness among currently pregnant women will usually underestimate the prevalence. Considering other characteristics of the condition, Christian (2000) has proposed eliciting the history of night blindness only from women whose last pregnancy ended in a live birth, and that this question be restricted to births in the last three years.

Because of health and survival risks associated with maternal night blindness, cases identified during antenatal clinics should be treated immediately (WHO recommendation is to treat with 10,000 IU/day or 25,000 IU/week for up to 3 months) or at least referred for treatment.

Maternal night blindness is now part of a core question in the DHS surveys. The data yield two indicators: per-

centage of women who report night blindness in the last pregnancy (in past three or five years) and an adjusted rate of percentage who report night blindness excluding those who report vision problems during the day. In low prevalence countries (less than five percent), because finding a widely recognized local term is difficult, the interviewers must be carefully trained to adequately describe the condition. The low prevalence of night blindness may necessitate large sample sizes to detect changes at the population level.

Recent DHS data from several African countries identified prevalences of night blindness that were lower than expected based upon other indicators of vitamin A status in those populations. The five percent cut-off point for maternal night blindness used to indicate vitamin A deficiency in a population was based largely on data from Asia. At the time of writing, further work is underway to confirm the appropriateness of this cut-off point and the adjustment process for this indicator in Africa.

Part III.G
Breastfeeding

- Timely initiation of breastfeeding: percent of infants 0 - < 12 months who were put to the breast within one hour of delivery

- Exclusive breastfeeding rate (EBR): percent of infants 0 - < 6 months of age who are exclusively breastfed

- Timely complementary feeding rate: percent of infants 6 - < 10 months given breastmilk and solid and/or semi-solid foods

- Lactational amenorrhea method acceptor rate (LAR): percent of eligible women who use LAM as their method of family planning

- Lactational amenorrhea method user rate (LUR): percent of women of reproductive age who use LAM as their method of family planning

In recent years, scientific knowledge has greatly advanced on the benefits of breastfeeding, the physiologic mechanisms underlying these benefits, and the optimal practice of breastfeeding. Epidemiologic research has clearly demonstrated that breastfeeding provides advantages to infants with regard to general health, growth, and development, while significantly decreasing the risk for a large number of acute and chronic diseases. Exclusively breastfed infants are at a much lower risk of infection from diarrhea and acute respiratory infections – two leading causes of infant death – than are infants who receive foods in addition to breastmilk during their first months of life. Breastmilk stimulates their immune systems and response to vaccinations, and it contains hundreds of health-enhancing antibodies and enzymes. Breastfeeding is the ideal method of feeding and nurturing infants. A final benefit of exclusive breastfeeding is that it protects the mother from pregnancy up to six months after delivery, if she does not resume menstruation.

Although the benefits of breastfeeding in terms of child survival are well known, changes in child mortality are difficult to measure and cannot easily be attributed to specific program interventions. Attitudes towards breastfeeding, awareness of the importance of exclusive breastfeeding, and support to enable mothers to breastfeed are important outcomes of promotional activities in health programs, but they may also be difficult to measure and/or interpret, and may not reflect actual practice. By contrast, indicators of **current** breastfeeding practices are relatively easy to measure and are sensitive to changes resulting from program activities.

Exclusive Breastfeeding from 0-<6 Months Followed by Complementary Feeding

Appropriate infant feeding practices are of fundamental importance for the survival, growth, development, health, and nutrition of infants and children everywhere. As such, the optimal duration of exclusive breastfeeding is a public health issue. Whereas consensus exists on the need for exclusive breastfeeding, debate continues on its optimal duration.

Early in 2000, WHO commissioned a systematic review of the published scientific literature on the optimal duration of exclusive breastfeeding; more than 3000 references were identified for independent review and evaluation. The process consisted of a global peer review, followed by a technical review by a group of breastfeeding experts held in Geneva (May 28-30, 2001). The results of this review were reported to the Fifty-fourth World Health Assembly in May 2001. The Assembly recommended support for exclusive breastfeeding for six months, followed by the introduction of complementary foods and the continuation of breastfeeding.

In developing countries, the most important advantage of exclusive breastfeeding for six months – as compared with exclusive breastfeeding for four months followed by partial breastfeeding to six months – relates to infectious disease morbidity and mortality, especially that due to gastrointestinal infection (diarrheal disease). The high incidence of, and mortality from, gastrointestinal infection suggests that exclusive breastfeeding for six months has a strong protective effect against diarrheal morbidity and mortality.

USAID, UNICEF, and WHO have all endorsed the recommendation for the introduction of nutritionally adequate, safe, and appropriate complementary foods, in conjunction with continued breastfeeding, at about six months of age. Breastmilk continues to be an important source of energy and protein, and other nutrients such as vitamin A, once complementary foods are introduced into the infants' diet, and breastmilk continues to protect the infant against disease during the second year of life.

Complementary feeding (CF) is a highly complex issue and includes factors such as density, quantity, quality, frequency, timeliness, food hygiene, as well as feeding during/after illness and care/active feeding. Inappropri-

ate complementary feeding practices, compounded by nutritionally inadequate and frequently contaminated foods often introduced too early or too late, remain a major cause of malnutrition.

Methodological Challenges of Evaluating Infant Feeding Programs

- **The use of 24-hour recall data tends to overestimate the percentage of infants who have been exclusively breastfed since birth.**

A 24-hour recall measure reflects current breastfeeding status and may cause the proportion of exclusively breast-fed infants to be slightly overestimated, since some infants who consume other liquids irregularly may not have received them in the 24 hours before the survey.

WHO's *Indicators for Assessing Breast-feeding Practices*, Wellstart International's *Tool Kit for Monitoring and Evaluating Breastfeeding Practices and Programs,* and the DHS reports all calculate the exclusive breastfeeding rate (EBR) using the 24-hour recall method. Using cross-sectional surveys, one can obtain the best estimates of exclusive breastfeeding from current status data that include all births within a specified time period. The advantage of this approach is that it is not subject to recall error. Evaluators should then interpret the measure as the percentage of infants who "are currently being exclusively breastfed" rather than the percentage that have been exclusively breastfed since birth.

- **Evaluators need large sample sizes to detect change in breastfeeding practices, but infants represent a small proportion of the population.**

Any assessment of behavioral change in infant feeding requires attention to the size of the comparison groups. The sample size depends on both the magnitude of the change and on the prevalence of the condition or practice. The detection of relatively small changes (e.g., five to ten percentage points) over time in breastfeeding and other infant feeding behaviors requires large sample sizes.

By contrast, simple monitoring of infant feeding practices does not require a specific sample size and can be very useful in tracking ongoing project outreach. However, monitoring neither allows for a rigorous evalua-

tion of change, nor measures actual prevalence of this behavior because of the small, nonrepresentative samples.

- **Infant feeding behavior data relies upon accurate age data of the infant.**

Although evaluators may track many health interventions with only a general reference to the child's age (e.g., less than one year), tracking breastfeeding practices requires accurate assessment of the infant's age. Interviewers can ascertain the age by first asking the mother for the infant's birthdate and then by confirming the birthdate with a child health card or other official registry of the child's birthdate.

- **Breastfeeding questions typically require more than a "yes" or "no" response.**

Multiple factors define whether breastfeeding is optimal, including the exact liquids and foods, if any, given in the preceding 24-hours. Ideally this list of liquids and foods will be comparable to the DHS with additional items that reflect local food preferences and food availability.

The data needed to calculate infant feeding behaviors related to exclusive breastfeeding 0-<6 months of age and timely complementary feeding 6-<10 months of age require that the interviewers ask the respondent about a series of foods given within the previous 24-hours. This line of questioning requires more than a "yes" or "no" response, thus reducing interviewer or respondent error. Interviewers should undergo intensive training on this set of items.

- **The accepted, standard complementary feeding indicator reflects general dietary intake of solid and semi-solid foods during a specified time period only.**

Complementary feeding, a highly complex issue, involves factors such as the quantity and quality of food, frequency and timeliness of feeding, food hygiene, and feeding during/after illness. Program personnel at the country level must consider these many factors when they try to address the problems of infant and young child feeding in the local context.

The standard CF indicator fails to account for program-specific or context-specific feeding recommendations

regarding the frequency, quality, or quantity of foods given during the proceeding 24 hours.

Two Age-groups for Optimal Infant Feeding

Two main types of indicators relate to optimal infant feeding: (1) those concerning breastfeeding behaviors during the first six months of life, and (2) those referring to the introduction of complementary foods while maintaining breastfeeding beginning at six months. These age groups reflect expert consensus as to the optimal time period for exclusive breastfeeding, as well as for the introduction of complementary foods to an infant's diet.

In population-based surveys, measuring these infant-feeding indicators requires sampling of infants 0-<6 months of age and infants 6-<12 months of age. Together, these groups represent the continuum of infant nutrition care in the first year of life. Used together, these data sets reflect the prevalence of optimal infant-feeding behaviors during the first year of life in a given population.

The main purpose of a common set of breastfeeding indicators is to standardize the assessment and evaluation of breastfeeding behaviors across programs implemented and funded by different organizations. The set of indicators in this section are limited in number, fairly easy to measure and interpret, and operationally useful. Moreover, they have been field-tested, are consistent with worldwide breastfeeding goals, and can be obtained from available DHS data.

Evaluators can use the indicators in this section as the outcome variables in measuring behavior change due to program interventions in the context of an experimental or quasi-experimental design. Evaluators can also calculate these indicators from program statistics for the purpose of tracking breastfeeding behavior among clients, but not for establishing the impact of a specific program or intervention on behavior of women with infants in the population in question.

TIMELY INITIATION OF BREASTFEEDING: PERCENT OF INFANTS 0 - < 12 MONTHS WHO WERE PUT TO THE BREAST WITHIN ONE HOUR OF DELIVERY

Definition
Timely initiation of breastfeeding is calculated as:

$$\frac{\text{\# of infants } 0 < 12 \text{ months put to the breast within 1 hour of delivery}}{\text{Total \# of infants } 0 < 12 \text{ months}} \times 100$$

Data Requirements
Recall data from mothers with infants less than 12 months old

Data Source(s)
Population-based surveys employing representative samples. The DHS reports the initiation of breastfeeding within one hour for those countries in which the breastfeeding/infant-feeding module is included in the DHS.

Evaluators may use program records to track trends in breastfeeding initiation among clients, but not to measure the impact of program interventions on women with infants in the population of the catchment area.

Purpose and Issues
Mothers are more likely to successfully initiate lactation, to encounter fewer problems breastfeeding, and to maintain optimal breastfeeding behaviors if they initiate breastfeeding shortly after birth.

Breastfeeding should begin no later than one hour after the delivery of the infant. Colostrum, the thick yellowish milk produced in the first few days after birth, is nutritious and helps to protect the infant against common infections. Thus, breastmilk is the infant's first "immunization" against common illnesses.

A mother may have difficulty correctly recalling (months after the event) when she initiated breastfeeding for her youngest infant; thus, this indicator is subject to potential recall bias. This bias is likely to be even greater in populations unaccustomed to remembering and conceptualizing time. However, because this particular type of bias (toward a longer or shorter period than actually occurred) is assumed to be randomly distributed across a population, the potential bias should not skew the data to misrepresent the population's general behavior related to breastfeeding initiation.

EXCLUSIVE BREASTFEEDING RATE (EBR): PERCENT OF INFANTS 0 - < 6 MONTHS OF AGE WHO ARE EXCLUSIVELY BREASTFED

Definition
Exclusive breastfeeding is the practice of giving only breastmilk to the infant, with no other solids or liquids, including water

Infants, are, however, allowed to have drops of vitamins/minerals/medicines.[1]

This indicator is calculated as:

$$\frac{\text{\# of infants 0 - < 6 months exclusively breastfed}}{\text{Total \# of infants 0 - < 6 months}} \times 100$$

Data Requirements
A 24-hour recall of food consumption of infants less than six months of age

Data Source(s)
Population-based surveys employing representative samples (e.g., the DHS) and program records of EBR (to track trends but not impact)

The DHS country reports and Nutrition Reports both present the EBR for infants 0-<4 months of age. However, evaluators can calculate the EBR for infants 0-<6 months using DHS data.

Purpose and Issues
Even in hot, dry climates, breastmilk contains sufficient water for an infant's needs. Additional water or sugary drinks are unnecessary to quench the infant's thirst, and they can also be harmful. If the infant is also given water, or drinks made with water, then the risk of diarrhea and other illnesses increases.

Although the benefits of breastfeeding in terms of child survival are well known, the effects of breastfeeding on child mortality are difficult to measure. Indicators of **current** breastfeeding practices, however, can be relatively easily measured and are sensitive to changes resulting from program activities.

Using a 24-hour recall period to measure current status may slightly overestimate the proportion of exclusively breast-fed infants because some infants who are given other liquids irregularly may not have received them in the 24 hours before the survey. WHO's *Indicators for Assessing Breast-feeding Practices*, Wellstart International's *Tool Kit for Monitoring and Evaluating Breastfeeding Practices and Programs*, and the DHS reports all calculate EBR using the 24-hour recall method.

Evaluators can obtain the best estimates of exclusive breastfeeding from current status data in cross-section surveys. The advantage of this approach is that it avoids subject recall error. Evaluators should interpret the measure as the percentage of infants who "are currently being exclusively breastfed," rather than the percentage exclusively breastfed since birth.

Gender Implications of this Indicator

The rate of exclusive breastfeeding, if disaggregated by sex, can be an indication of whether gender bias exists in the country. In India, women more often discontinue breastfeeding of daughters in the first six months as compared to sons. A nutritional study of weight for age among boys and girls demonstrates how "broad nutritional symmetry (at birth) between boys and girls turns into a situation of significant female disadvantage" (Sen, 2001). Discontinuation of exclusive breastfeeding is one of several factors ultimately contributing to a lower female/male sex ratio in India as compared to countries where son preference is not evident.

---

[1] This is the WHO definition of exclusive breastfeeding, 1991c, adopted thereafter by international agencies, including USAID.

TIMELY COMPLEMENTARY FEEDING RATE: PERCENT OF INFANTS 6 - < 10 MONTHS GIVEN BREASTMILK AND SOLID AND/OR SEMI-SOLID FOODS

Definition

Complementary foods are defined as solid or semi-solid/mushy foods; complementary foods do **not** include fluids.

This rate can be calculated as follows:

$$\frac{\text{\# of infants 6 - < 10 months who have received solid and/or semi-solid foods}}{\text{Total \# of infants 6 - < 10 months}} \times 100$$

Data Requirements

A 24-hour recall of food consumption of infants 6 - <10 months of age

Data Source(s)

Population-based surveys employing representative samples (e.g., the DHS). Evaluators may use program records to track trends in complementary feeding but not to measure impact. DHS reports present data for this indicator for those countries in which the breastfeeding/infant-feeding module was included.

Purpose and Issues

This basic calculation of complementary feeding uses 24-hour recall. Evaluators may supplement these data by an additional indicator(s) reflecting program messages relating to quantity, density, and/or quality of complementary foods. By the age of six months, almost all infants should receive complementary foods in addition to breastmilk.

This indicator has several limitations. First, it reflects only the consumption of complementary feeding, not the appropriateness of those foods. Second, it provides minimal information on the extent to which infants are fed according to prescribed guidelines.

If researchers or evaluators opt to collect additional information on complementary feeding (e.g., for the purpose of evaluating a specific program intervention), we recommend retaining this "basic" indicator as well, for comparisons with other populations.

LACTATIONAL AMENORRHEA METHOD ACCEPTOR RATE (LAR): PERCENT OF ELIGIBLE WOMEN WHO USE LAM AS THEIR METHOD OF FAMILY PLANNING

Definition

The percent of women giving birth in a reference period who consciously and deliberately accept LAM as a modern contraceptive method

This indicator is calculated as:

> # of women using LAM as an FP method   x 100
> Total # of women with infants < 6 months

Data Requirements

Total number of women with infants less than 6 months old, and of those, the number who choose to use LAM as a method of family planning

Data Source(s)

Population-based surveys employing representative samples (e.g., DHS), or family planning service statistics (if data are systematically obtained on the age of the youngest child)

Purpose and Issues

The Lactational Amenorrhea Method (LAM) is a natural family planning method that protects a woman from pregnancy by suppressing ovulation during the first six months after delivery, provided that she meets three criteria:

1. The woman has not resumed her menstrual period; **and**

2. The infant is fully or nearly fully breastfed;[2] **and**

3. The infant is less than six months old.

When any one of these three criteria is no longer met, another family planning method must be introduced quickly to ensure birth spacing. Evaluators can capture this follow-up method in both the family planning registry as well as in subsequent population-based surveys.

One shortcoming of this indicator is that it is often based on self-report, without verification that the respondent actually fulfills the three criteria listed above. A more precise measure will include questions that confirm that the respondent knows the three criteria and that she meets them.

---

[2] Full or nearly full breastfeeding significantly impacts fertility. This category includes exclusive, almost exclusive and high (partial) breastfeeding. Thus, the infant can receive only breastmilk or mostly breastmilk with some addition of liquids such as juice or ritualistic foods given infrequently.

LACTATIONAL AMENORRHEA METHOD USER RATE (LUR): PERCENT OF WOMEN OF REPRODUCTIVE AGE WHO USE LAM AS THEIR METHOD OF FAMILY PLANNING

Definition

The percent of women of reproductive age using the lactational amenorrhea method (LAM) as a modern contraceptive method, at a given point in time (e.g., at the time of the survey)

This indicator is calculated as:

$$\frac{\text{\# of married women of reproductive age using LAM as an FP method}}{\text{Total \# of married women of reproductive age}} \times 100$$

Data Requirements

The total number of married women of reproductive age, and of these, the number who choose LAM as their method of family planning

Data Source(s)

Population-based surveys employing representative samples (e.g., the DHS); or family planning service statistics

Purpose and Issues

This indicator measures the percentage of married women of reproductive age in a given population using the LAM method. As such, it reflects the use of LAM relative to other family planning methods. In fact, the comparison is somewhat misleading, in that LAM can only be practiced by women with a child less than six months, whereas other methods are potentially available for use to all women of reproductive age (medical contraindications being the exception). Ideally, evaluators will measure the LUR from a population-based representative survey. Service statistics also constitute a source for calculating LUR, but the findings will not be generalizable to the larger population.

The LUR (based on all married women of reproductive age) will be lower than the LAR (based on women with an infant less than six months). Both are useful, but for different purposes. The LUR measures the use of LAM in relation to other contraceptive methods. The LAR reflects the extent of use among **potential** users (i.e., women with a child less than six months old).

- Existence of supportive ARH policies

- Adolescents are meaningfully involved in the design and implementation of the program

- Percent of program staff trained to work with or provide services to adolescents

- Percent of adolescents aware of the program

- Youth friendliness of reproductive health services

- Sexual-reproductive health education curriculum conformity to "best practices"

- Percent of adults in community who have a favorable view of the program

- Number/percent and characteristics of adolescents "served" or "reached" by the program

- Sexual-Reproductive Health (SRH) knowledge

- Percent of adolescents who have "positive" attitudes toward key sexual-reproductive health issues

- Percent of adolescents who are confident that they could refuse sex if they didn't want it

- Percent of adolescents who are confident that they could get their partner(s) to use contraceptives/condoms if desired

- Percent of adolescents who have ever had sexual intercourse

- Age at first intercourse

- Number of sexual partners among sexually active adolescents during a specified reference period

- Number/percent of adolescents who have experienced coercive or forced sex

- Percent of sexually initiated adolescents who used a condom at first/last sex

- Percent of sexually active, unmarried adolescents who consistently use condoms

- Percent of adolescents who regularly use drugs/alcohol

- Percent of adolescents who feel "connected" with their parents/family

- Percent of adolescents who have ever been pregnant or caused a pregnancy

- Unmet need for family planning among adolescents

- Percent of adolescents who were ever diagnosed with an STI

ADOLESCENT REPRODUCTIVE HEALTH PROGRAMS

Recent years have witnessed a dramatic increase in the attention and resources directed to the reproductive health of adolescents (10- to 19-year-olds) and young adults (20- to 24-year-olds) on more or less a global basis. Among the reasons for this increase are (1) general concerns over the health and human rights of youth, (2) the demographic significance of the 10-19 age group in many developing countries, and (3) the pivotal role adolescents and young adults will play in the HIV/AIDS epidemic.

Adolescent reproductive health (ARH) programs focus on achieving one or more of four major goals: (1) creating an enabling and supportive environment for youth, (2) improving the knowledge, attitudes, skills, and behaviors of adolescents, (3) increasing adolescents' use of services, and (4) increasing adolescents' participation in programs. Programs may assume a variety of forms and may appear in a variety of settings. Among the more common program types are:

- Sexual-reproductive health or life-skills education programs in schools;
- Mass media-based behavior change and social marketing interventions;
- Programs to make reproductive health services more "youth friendly;"
- Community-based non-formal education programs;
- Workplace-based reproductive health education programs;
- Youth clubs/organizations;
- Livelihood programs to generate economic opportunities for youth;
- Advocacy campaigns to influence political and cultural leaders (and adults in general); and
- Community mobilization campaigns.

Surveys of the intended audience for adolescent programs are a primary vehicle for collecting data for evaluation purposes. The intended population for such surveys will vary depending upon the group of youth the program intends to reach (e.g., youth attending school, youth employed in the informal sector of the economy, all youth).

Methodological Challenges of Evaluating Adolescent RH Programs

As they do in other areas of reproductive health, program officials and evaluators face a number of formidable methodological challenges in assessing the performance of adolescent reproductive health (ARH) programs. Together, these challenges make the evaluation of ARH programs among the more difficult types of reproductive health programs.

Specific methodological challenges to evaluating ARH programs include the following:

- **A myriad of factors heavily influence adolescent behaviors.**

Adolescent behaviors are influenced in important ways by a sizeable number of factors operating at the individual, family, school, community, and societal levels. Granted, these same factors influence adults. However, because adolescents have not fully developed – socially, psychologically, and physically they are perhaps more susceptible to "contextual" or "environmental" influences than are adults. This susceptibility requires that programs address a number of determinants or "antecedents" of adolescent behaviors simultaneously. Evaluators must measure and "control for" a sizeable number of factors in order to tease out the effects of specific ARH interventions. Furthermore, evaluators often find themselves beyond the bounds of their own disciplinary training in dealing with the range of factors (e.g., relationships with family, school, and community; self-esteem; self-efficacy).

- **The intended effects of ARH interventions are long-term for some interventions, further complicating evaluation.**

The appropriate time-reference for measuring the impact of ARH programs is tricky. For some outcomes, (e.g., delayed age of sexual initiation), the desired result/behavior is a short-term phenomenon an evaluator can accurately measure within the typical time-frames of most ARH program evaluations (usually two to three years or less). For other outcomes, however, evaluators require longer periods of observation. For example, life-skills education programs help youth develop skills that will be manifest over their adolescent and adult years. Few evaluators have assessed the long-term effects of such interventions; the best they often can do is to measure impact over a relatively short period of time.

Further complicating matters is that, in some cases, program effects may be short-term or transitory in nature. For example, an evaluation of school-based ARH-education programs in Jamaica found significant effects on knowledge, attitudes and behaviors when measured nine months after program implementation, but these effects had largely disappeared when measured again after 21 months (Eggleston et al., 2000). Thus, strong impact evaluations of ARH programs require evaluators to measure impact at several points in time after program implementation.

- **Measuring the quality of ARH programs requires an understanding of cultural constructs in the local setting.**

Assessing the quality of ARH programs from the "client's perspective" requires the evaluator to elicit subjective interpretations, perspectives, and meanings from youth and others in the community. These elements are critical in designing effective programs and in appropriately evaluating them. As a result, a combination of qualitative and quantitative data are generally required for the meaningful evaluation of ARH programs.

- **ARH programs are often quite complex, multi-component initiatives.**

Because ARH programs must simultaneously address multiple "risk" and "protective" factors, a sizeable and growing number of programs have complex designs and multiple components. For example, many programs have life-skills education (both formal and non-formal), peer promotion, community mobilization, and access to RH services components. The nature of such programs dictates that evaluators measure multiple processes and outcomes to evaluate the program. Measuring the impact of the separate components of such programs is especially difficult, and as a result, program evaluations often focus on the net or combined impact of the full "package" of interventions.

- **ARH programs produce effects at more than one level.**

Although ARH programs primarily focus on influencing adolescent behaviors and RH outcomes, programs often attempt to bring about change at more than one level. For example, some programs mobilize community support for and involvement in initiatives and activities for youth. Failure to garner such community involvement could greatly diminish the effectiveness of the program in changing the attitudes and behaviors of adolescents at the individual level. Without measuring change (or lack thereof) at the community level, the evaluator could not accurately interpret the lack of change at the individual level.

- **Sensitivities to ARH programs and to issues of adolescent sexuality complicate measurement in many settings.**

Many societies regard the intended outcomes of ARH programs as personal and private. Some societies even prohibit discussions about sexual behavior and personal relationships. Program officials and evaluators may also face parental and community resistance to asking adolescents questions about these topics. Community leaders and other stakeholders may believe that the young people in their communities do not engage in risky behaviors, and therefore evaluators do not need to ask questions about these topics. Stakeholders may also fear social or political danger in uncovering the truth about young people's behaviors and may attempt to block data collection. Because of the social sensitivities surrounding adolescent sexual behaviors, evaluators face more rigid informed/parental consent procedures for ARH programs than for other types of RH programs.

- **Overlap of Indicators with Other Areas of Reproductive Health**

Some indicators described elsewhere in this *Compendium* are relevant to ARH programs. For example, most or all indicators pertaining to policy, program management, commodities and logistics, management, BCC,

and training are generic indicators that apply to ARH programs as well as to other types of RH programs. However, the nuances involved in ARH programs necessitate several specific indicators even in these generic areas. For example, ARH programs emphasize youth participation in program design and implementation, on specific characteristics of health facilities that attract youth, and specialized training of program staff to serve and to work with youth. In terms of outcomes, whereas ARH programs share many intended outcomes with family planning and with other RH programs (e.g., increased contraceptive use, reduced rate of unwanted pregnancies, reduced rate of STIs), other outcomes are unique to ARH programs. For example, because ARH programs often aspire to influence the broad social environment in which youth are reared, outcome indicators pertaining to matters such as community support for programs and services directed to youth and to "connections" with parents and family are relevant both as program outcomes and as contextual factors shaping program design. ARH programs also often attempt to reduce the prevalence of non-sexual risk behaviors (e.g., alcohol and drug use) and to develop specific skills and competencies ("self-efficacy") as key intermediate outcomes that improve the reproductive health of adolescents. It should be noted that this section describes the key specialized indicators for ARH programs. Finally, evaluators should collect and report outcome indicators for adolescent programs by gender.

EXISTENCE OF SUPPORTIVE ARH POLICIES

Definition

This indicator is a composite index measuring the extent to which the overall policy environment in a country supports adolescent reproductive health concerns. The index assesses the existence of:

- Policy or legislation recognizing the rights of adolescents, including unmarried adolescents, to receive reproductive health services;
- A formal policy setting a minimum age for marriage;
- Policies prohibiting sexual exploitation and/or violence;
- Policy or legislation authorizing sexual-reproductive health education in schools (or lack of restrictive policies or legislation);
- Permission for pregnant adolescents to continue their education;
- Policy or legislation authorizing sales of contraceptives to youth in both the public and commercial sectors;
- Public health sector service delivery guidelines mandating the provision of all reproductive health services to all adolescents; and
- Government authorization of media campaigns on ARH issues.

Evaluators score the index by assigning a value of 2 when the policy environment fully satisfies a given condition, 1 when it partially satisfies the condition, and 0 when it fails to satisfy the condition.

Data Requirements

Evidence of the presence or absence of each of the items included in the index

Data Source(s)

Government documents or other means of verifying the existence of relevant policies, legislation, or regulations; interviews with government officials and key informants

Purpose and Issues

As is the case for family planning programs, a favorable (or at least not hostile) policy environment is essential for the operation and expansion of adolescent reproductive health programs. Although few countries have explicitly restrictive policies, because of the social sensitivity surrounding ARH issues, many countries lack formal policies regarding the provision of RH information and services to youth. The absence of formal policies permits administrators and service providers to impose restrictions – based on their personal beliefs – that prohibit youth from gaining access to essential information and services. This indicator measures whether formal policies that enable and support the provision of RH information and services to youth have been enacted. However, because educators and service delivery staff have personal biases toward and discomfort in addressing adolescent sexual-reproductive health issues, the mere existence of policies does not guarantee the implementation of those policies. Evaluators may expand the indicator to include scores on the extent to which each policy is actually being implemented. A separate indicator measures the actual availability of and access to relevant information and services by youth.

ADOLESCENTS ARE MEANINGFULLY INVOLVED IN THE DESIGN AND IMPLEMENTATION OF THE PROGRAM

Definition

This qualitative (yes/no) indicator measures adolescent participation in a program. The evaluator assigns a "yes" score if adolescents participated in the program in a meaningful way. Two areas for participation include: (1) program design (did adolescents from the intended audience participate in designing the program by communicating their needs and preferences) and (2) program implementation (did the adolescents help implement the program).

Data Requirements

Program documents or other evidence that (1) the program designers assessed the needs of the program's intended audience through a participatory process entailing significant input from youth in the program's intended audience, (2) the findings from the assessment helped shape program design and strategy development, and (3) youth play key roles in program management or in the delivery of services.

Data Source(s)

Program records; interviews with program staff; interviews with adolescents participating in the assessment and program design; interviews with youth involved in program implementation

Purpose and Issues

Most ARH experts concur that youth participation in program design and implementation enhances program appeal and effectiveness. This indicator provides a qualitative measure of the extent of meaningful participation by youth in the program's design and implementation. Youth have "meaningful participation" if they play a major role in carrying out the assessment, in deriving conclusions from the assessment data gathered, in designing the program, and in managing and carrying out program activities.

PERCENT OF PROGRAM STAFF TRAINED TO WORK WITH OR PROVIDE SERVICES TO ADOLESCENTS

Definition

The percent of program staff specifically trained to work with or provide information, education, or family planning services to adolescents

This indicator is calculated as:

$$\frac{\text{\# of program staff who have received specific training to provide education/counseling or adolescent health care}}{\text{Total \# of program staff working with adolescents}} \times 100$$

Data Requirements

Number of program staff working with adolescents, number (of these) who received specific training to provide education/
counseling or adolescent health care

Data Source(s)

Program personnel files/records

Purpose and Issues

Working with youth requires perspectives and skills often lacking in standard pre-service training. This indicator measures the extent to which program personnel working with adolescents have received specific training to provide services to adolescents. Services may include outreach, information, education, counseling, referral, and reproductive health services. Note that this indicator only measures staff exposure to training; it does not measure the quality of the training or the staff competence in working with adolescents as a result of the training.

PERCENT OF ADOLESCENTS AWARE OF THE PROGRAM

Definition
The percent of adolescents who report knowing of the program's services and/or activities

The services and activities will be specific to each program. Thus, the indicator may refer to sexual-reproductive health or life-skills education in schools or workplaces; reproductive health services at clinics or youth centers; the existence of youth organizations, and radio or television programs for youth.

This indicator is calculated as:

$$\frac{\text{\# of adolescents aware of the program}}{\text{Total \# of adolescents}} \times 100$$

Data Requirements
Responses to survey questions on awareness of the program's existence and activities by adolescents. The preferred procedure is to first ask about program's services for youth without prompting; then, for adolescents who do not spontaneously report knowledge of the program, to identify the program and ask if the respondent has heard of it.

Data Source(s)
Survey of adolescents in the program's intended population

Purpose and Issues
Before adolescents can use a program, they must know it exists. This indicator provides program managers with a basis for assessing whether they must initiate promotional or awareness-raising activities as part of their youth initiative.

Youth Friendliness of Reproductive Health Services

Definition

This indicator is a composite index measuring whether reproductive health services are "youth friendly." Services are "youth friendly" if they "have policies and attributes that attract adolescents to the facility or program, provide a comfortable and appropriate setting for youth, meet the needs of adolescents, and are able to retain their adolescents for follow-up and repeat visits" (Senderowitz, 1999). Aspects of an "adolescent friendly" environment can include space or rooms dedicated to ARH services, policies and procedures to ensure privacy and confidentiality, peer educators on site, nonjudgmental staff, and acceptance of drop-in clients.

Data Requirements

Evidence as to whether reproductive health services satisfy standards for being "youth friendly." The following characteristics make facilities/services "youth friendly:"

- Facility hours are convenient for adolescents;
- Facility location is convenient for adolescents;
- Facility has adequate space and sufficient privacy;
- Facility has comfortable surroundings for adolescents;
- Staff have been specially trained to work with or to provide services to adolescents;
- Staff treat adolescent clients with respect;
- Staff honor privacy and confidentiality of adolescent clients;
- Staff allow adequate time for adolescent client and provider interaction;
- Peer counselors are available on site;
- Young male clients are equally welcomed and served as young female clients are;
- Group health discussions are available;
- Necessary referrals are available;
- Service fees are affordable for adolescent clients;
- Drop-in clients are welcomed, and appointments are arranged rapidly;
- A wide range of reproductive services are available;

- Educational materials are available on site for clients to take home;
- Adolescents perceive that they are welcome regardless of their age and marital status; and
- Adolescents perceive that providers will be attentive to their needs.

Evaluators create this index by assigning a score to each item: 2 points for complete fulfillment of the condition, 1 point for partial fulfillment of the condition, and 0 for lack of fulfillment. Evaluators may derive a total facility score if they first sum the item scores and then divide that result by the total number of points possible (Nelson, MacLaren, and Magnani, 2000).

Data Source(s)

Facility records; facility inventories; interviews with adolescent clients, providers, and managers at clinics; client exit interviews; interviews of youth in the community

Purpose and Issues

Because reproductive health services in most settings have been designed for older, married women, unmarried female and male adolescents face a variety of barriers to service use. Among these are policies that restrict their access to services and information, negative community attitudes toward providing reproductive health services to unmarried adolescents, adolescent embarrassment at being seen at facilities, and fear that the facility will not honor privacy and confidentiality.

To overcome these barriers, a number of service-providing organizations have sought to make their services more "youth friendly." By offering more youth-friendly reproductive health services, programs may effectively attract young people and may provide quality reproductive health services in a comfortable and responsive environment. Adolescents can receive services in a health facility, such as a clinic, health post or hospital, from trained personnel who provide services in a workplace or school setting, through community outreach workers or peer educators. Regardless of the venue,

services must have special characteristics that attract, serve, and retain adolescent clients.

This indicator is most appropriate for assessing facilities and services that were not specifically designed for adolescents (such as a family planning clinics, health posts, or pharmacies), because adolescent facilities were presumably designed with the characteristics of adolescent friendliness in mind. However, this indicator can also monitor the adolescent friendliness of adolescent-centered facilities over time. For example, after a baseline assessment, the program manager may plan to make changes in services over the next 6 months and may allow those changes to become part of the service-delivery protocols over the next 12 months. The program manager may then decide to undertake a follow-up assessment 18 months later to determine if the changes occurred. The follow-up assessment should measure the same characteristics it measured in the initial assessment.

SEXUAL REPRODUCTIVE HEALTH EDUCATION CURRICULUM CONFORMITY TO "BEST PRACTICES"

Definition

This qualitative (yes/no) indicator measures the extent to which the program's sexual reproductive health (SRH) education curriculum contains all (or most) of the features identified as "best practices" or "key elements" of effective SRH programs. Alternatively, the indicator can serve as an index or scale indicating the percent of best practices and key elements that the program has incorporated into it's curricula and materials.

Data Requirements

Content analysis of the curriculum; accompanying materials; and activities that permit an assessment of conformity with "best practices"

Data Source(s)

Content analysis of program curriculum, materials, and learning methodologies; observation of actual delivery; interviews or focus groups with youth; or self-reported questionnaires from youth who participated in the program

Purpose and Issues

This indicator measures the quality of sexual-reproductive health education efforts focusing on curriculum content. The indicator reflects how well the program covers key aspects of sexual-reproductive health education and how appropriate the content is for the age-group of adolescents reached. Setting universally appropriate criteria is difficult because of cultural and socio-economic differences across and within countries. However, a growing consensus requires that such programs should cover, at minimum, the following: interpersonal communication, self-esteem, value clarification, life stage, decision-making, education and career goals, gender roles, dating, sexuality, marriage, and contraception. Illustrative guidelines for sexuality education in the U.S. – provided by SIECUS (1996) and

Kirby (2001) – enumerate ten characteristics that successful sexual-reproductive health education programs in the U.S. share:

- Focus on reducing one or more sexual behaviors that lead to unintended pregnancy or STI/HIV;
- Design based on theoretical approaches demonstrated to effectively influence health-related risky behaviors;
- Clear messages about sexual activity and condom/contraceptive use and continual reinforcement of the messages;
- Basic, accurate information about the risks of adolescent sexual activity and about methods of avoiding intercourse and using protection against pregnancy and STIs;
- Activities addressing social pressures that influence sexual behavior;
- Provides role modeling and practice communication, negotiation, and communication skills;
- Varied, participatory teaching methods that encourage participants to personalize the information;
- Incorporates behavioral goals, teaching methods, and materials that are appropriate to the age, sexual experience, and culture of the students;
- Sufficient duration to cover key topics and complete important activities; and
- Teachers and/or peer leaders who believe in the program and are adequately trained.

Program delivery may be non-didactic and thus more effectively reach adolescents; for example, seminars, drama events, musical presentations, sports.

These criteria, developed from programs in the United States, have yet to be validated for developing countries.

PERCENT OF ADULTS IN COMMUNITY WHO HAVE A FAVORABLE VIEW OF THE PROGRAM

Definition

The percent of adults from the intended audience in the geographic area covered by the program who report that they "like," "support," or "agree" with the goals, objectives, and activities of the program

This indicator is calculated as:

$$\frac{\text{\# of adults who have a favorable view of the program}}{\text{Total \# of adults}} \times 100$$

Data Requirements

Responses to survey questions on adult views of the program

Data Source(s)

Surveys of adults in the population covered by the program

Purpose and Issues

Although a positive image among adolescents is the most crucial, parental and adult perceptions of ARH programs are also important to program success in view of the key role adults play in shaping adolescent attitudes and perceptions. If parents and adults in the community disapprove of a program, their lack of support often influences the attitudes and behaviors of adolescents. The importance of adult perceptions and support are demonstrated in a recent study in Zambia, which found that trends in adolescent use of reproductive health services were more strongly associated with adult acceptance of providing such services to youth than attributes of the services themselves were (Nelson, Magnani, and Bond, 2001).

NUMBER/PERCENT AND CHARACTERISTICS OF ADOLESCENTS "SERVED" OR "REACHED" BY THE PROGRAM

Definition
The number of adolescents who have received program services, have participated in program activities, and have been exposed to program mass media messages

The evaluator can subdivide the total number exposed by the type of activity: school-based program, clinical services, youth center activity. In addition, surveys (if used) can show the percent reached by mass media messages. The evaluation can also classify participants in these activities by relevant characteristics such as: age, gender, marital status, race/ethnicity, socio-economic status, school matriculation status, employment status, pregnancy history, STI history, and contraceptive use history.

This indicator is calculated as:

$$\frac{\text{\# of adolescents served or exposed to the program}}{\text{Total \# of adolescents in the intended population}} \times 100$$

Data Requirements
Program service statistics or comparable data indicating the number and characteristics of adolescents served by the program; responses to survey questions on exposure to or participation in program activities. Where feasible, evaluators can collect comparable data on adolescents not served or reached by the program to verify that the program is reaching its intended audience and to identify under-served segments of the adolescent population.

Data Source(s)
Program records or surveys of the program's intended population/audience In programs that provide different types of services (e.g., youth centers offering recreational, educational, and health services), evaluators should compile service statistics separately for each major type of service or activity.

Purpose and Issues
This indicator measures the volume and characteristics of adolescent clients who participate in program activities or use program services. The exact wording of the indicator will vary by type of program.

Evaluators can readily compile data on the number and characteristics of adolescents that attend program activities or seek clinical services at fixed sites. To measure the reach of mass media and similar programs, evaluators can survey the intended audience and thus obtain counts or estimates of the percentage of adolescents "exposed" to specific communication programs.

In addition to the number and percent of adolescents in the intended population served or reached by the program, the evaluator should ascertain that the program reaches key sub-groups of adolescents. For example, health facility-based programs that reach primarily older, married females who have previously been pregnant will likely have a very different population impact than will comparable programs that reach younger, unmarried adolescents of both genders. Similarly, "low-risk" youth recruited as peer promoters who contact and engage other low-risk youth will likely have a very different population impact than will higher-risk youth recruited to contact other higher-risk youth. In short, the evaluator needs to verify that the program is reaching the sub-groups of interest within the population at large.

SEXUAL-REPRODUCTIVE HEALTH (SRH) KNOWLEDGE

Definition

This indicator is a composite indicator or index measuring adolescents' knowledge of key sexual-reproductive health (SRH) topics and issues. The topics and issues included in the indicator should reflect those of primary importance for protecting the reproductive health of adolescents and/or those the program emphasized.

Data Requirements

Evidence of knowledge of key SRH issues, usually solicited by means of personal interviews with or self-administered questionnaires completed by adolescents

The following is an illustrative list of topics that evaluators may include:

- The female menstrual cycle and conception;
- Ways to avoid pregnancy;
- Methods of contraception;
- Correct use of (at least) condoms and oral contraceptives;
- Existence of sexually transmitted infections (STIs);
- Means of transmission of STIs;
- Ways to avoid STIs; and
- Symptoms of STIs.

Data Source(s)

Surveys of adolescent program participants or of adolescents in the program's intended population

Purpose and Issues

Adolescents must have knowledge of key sexual-reproductive health topics and issues if they are to make informed decisions to protect their health and well being. Many adolescents get their RH information from poorly informed sources (i.e., peers). Inaccurate beliefs concerning levels of risk associated with particular behaviors and/or the effectiveness and side-effects of different types of contraceptives can be strong enough to prevent adolescents from accurately perceiving the potential consequences of their behaviors. This indicator is a composite measure that includes the SRH topics and issues of primary importance for protecting the reproductive health of adolescents and/or those topics and issues the program emphasized.

When interviewers question adolescents about these topics they should use local, non-scientific names to describe certain practices and conditions. Evaluators can and should analyze separately the individual topics and questions included in the composite index to determine those specific topics requiring further emphasis by the program. Although adolescents need accurate knowledge of SRH topics for informed decision-making, adolescents may not act in a manner consistent with their knowledge, such that evaluators need to measure behavior separately.

PERCENT OF ADOLESCENTS WHO HAVE "POSITIVE" ATTITUDES TOWARD KEY SEXUAL-REPRODUCTIVE HEALTH ISSUES

Definition

This composite indicator or index measures adolescents' attitudes toward key sexual-reproductive health (SRH) topics and issues. "Positive" attitudes are those logically expected to lead to positive RH outcomes. The topics and issues included in the indicator should reflect those of primary importance for protecting the reproductive health of adolescents and/or those the program emphasized.

This indicator is calculated as:

> # of adolescents who have positive
> attitudes toward key SRH issues _____ x 100
> Total # of adolescents

Data Requirements

Evidence of the prevalence of "positive" attitudes toward key SRH issues, usually solicited by means of personal interviews with or self-administered questionnaires completed by adolescents. Evaluators should tabulate data for this indicator by gender and age. The following items illustrate possible attitudes to measure:

Attitudes toward contraceptives/condoms:
- Condoms do/do not reduce sexual pleasure;
- Carrying condoms is/is not difficult;
- Using condoms is/is not a sign of mutual respect;
- Condoms are easy/difficult to obtain and use;
- My partner would/would not reject me if I insisted on condom use;
- Unmarried adolescents should/don't need to use condoms in all sexual encounters; and
- I am/am not responsible for my own well-being.

Gender-role stereotypes:
- Women who carry condoms are "easy" or prostitutes;
- Having sex with many women is a sign of manhood;
- "Real men" don't use condoms; and
- The female (sexual partner) is responsible for protection.

Attitudes toward abstinence:
- It is OK for youth to wait for marriage to have sex; and
- My friends would/would not laugh at me for refusing to have sex.

Perceived vulnerability:
- It (pregnancy/STIs) won't happen to me;
- Young people are healthy and don't need to worry about STIs; and
- Women can/cannot get pregnant the first time they have sex.

Data Source(s)

Surveys of adolescent program participants or of adolescents in the program's intended population

Purpose and Issues

Developing "positive" attitudes toward key sexual-reproductive health topics/issues is an important objective of many ARH programs. This indicator is a composite measure that covers attitudes toward the SRH topics and issues of primary importance for protecting the reproductive health of adolescents and/or those the program emphasized. As with knowledge, however, positive attitudes do not necessarily predict future behaviors.

PERCENT OF ADOLESCENTS WHO ARE CONFIDENT THAT THEY COULD REFUSE SEX IF THEY DIDN'T WANT IT

Definition

The percent of adolescents reporting confidence that they could refuse sex if they did not desire it

This indicator is calculated as:

> # of adolescents reporting that they could
> refuse sex if they did not desire it   x 100
> Total # of adolescents

Data Requirements

Responses to survey questions on whether adolescents are "confident," "somewhat confident," "unsure," or "not confident" that they could resist having sex when they did not desire it

Data Source(s)

Surveys of program clients/participants or adolescents in the program's intended population

Purpose and Issues

This indicator measures the level of confidence or "perceived self-efficacy" of adolescents to refuse sexual advances when they do not want to have sexual relations. A growing consensus claims that ARH education programs are most successful when they address social pressures that influence sexual behaviors. Many programs include exercises and "role-plays" on how to resist pressure tactics and to escape situations that may lead to sex, through negotiation and other tactics. Thus, the indicator can measure the effectiveness of such skill-based educational programs in increasing adolescents' self-efficacy with regard to resisting unwanted sexual pressures and advances. This indicator measures perceived self-efficacy, which may or may not correspond to actual responses to real-life situations.

Because responses to sexual advances are likely to be context specific, the preferred measurement approach is to solicit responses to various situations that adolescents might find themselves in. For example, the interviewer may ask respondents how confident they are in their ability to refuse sex with:

- A person they have known for days;
- A person they have known for months;
- A person who offers them gifts;
- A person whom they care about deeply;
- A person who has paid for their school or training fees and who demands sex; and
- A person who has power over them, such as a teacher or an employer.

Gender Implications of this Indicator

Recent surveys indicate that for girls, the first experience of sexual intercourse is often involuntary (in some but not all developing countries). Forced sex (rape) is a form of gender-based violence. Many girls are coerced into sex by older men who view younger partners as less likely to have an STI. Some men believe that sex with a virgin can cure them of HIV/AIDS. Young girls say they lack the skills and self-confidence to refuse a more powerful and older male. Economic realities for many young girls makes refusing sex difficult and increases the likelihood that they will trade sex for money or gifts. In sub-Saharan Africa and other countries, these factors have led to new HIV infections among adolescent girls that are higher than those among boys and adults of either sex.

Percent of Adolescents Who are Confident That They Could Get Their Partner(s) to Use Contraceptives/Condoms if They Desired

Definition

The confidence or "self-efficacy" of adolescents in their ability to negotiate contraceptive/condom use with their partner(s)

This indicator is calculated as:

$$\frac{\text{\# of adolescents reporting ability to negotiate contraceptive/condom use with their partners}}{\text{Total \# of adolescents}} \times 100$$

Data Requirements

Responses to survey questions on whether adolescents are "confident," "somewhat confident," "unsure," or "not confident" that they could convince their partner(s) to use a contraceptive/condom if desired

Data Source(s)

Surveys of program clients and participants or adolescents in the program's intended population

Purpose and Issues

This indicator measures the level of confidence or "perceived self-efficacy" of adolescents – desiring protection – to successfully negotiate contraceptive/condom use with their partner(s) if desired. Like the ability/skill to resist social pressure to have sex, many ARH education programs emphasize negotiation skills with

regard to contraceptive/condom use. This indicator is particularly important for girls in developing countries, because many have limited negotiation skills or power to convince sexual partners to use contraceptives/condoms. The indicator can thus measure the effectiveness of such skill-based educational programs in increasing adolescents' self-efficacy at contraceptive use. Like the previous indicator, this indicator measures perceived self-efficacy, which may or may not correspond to actual behaviors in real-life situations.

Because self-efficacy of contraceptive use is also likely to be context specific, the preferred measurement approach is to solicit responses to various situations that adolescents may find themselves in. For example, an interviewer may ask respondents how confident they are in their ability to successfully negotiate contraceptive/condom use with:

- A person they have known for days;
- A person they have known for months;
- A person who offers them gifts;
- A person whom they care about deeply;
- A person who has paid for their school or training fees and who demands sex; and
- A person who has power over them, such as a teacher or an employer.

PERCENT OF ADOLESCENTS WHO HAVE EVER HAD SEXUAL INTERCOURSE

Definition

The percent of adolescents who have ever engaged in sexual intercourse (interpreted in most contexts to mean penile-vaginal intercourse)

This indicator is calculated as:

$$\frac{\text{\# of adolescents who have ever had sexual intercourse}}{\text{Total \# of adolescents}} \times 100$$

Data Requirements

Responses to a survey question asking whether they have ever had sexual intercourse

The question or questionnaire should specify penile-vaginal intercourse in order to minimize confusion as to the behavior the question referred to. The evaluators should measure the indicator for both married and unmarried youth.

Data Source(s)

Population-based surveys

Purpose and Issues

This indicator determines the extent to which adolescents in a program's intended population are sexually initiated. The indicator is useful both for designing ARH programs and for evaluating the effectiveness of existing programs aimed at postponing age at sexual debut. Because of large differences in age of the partners, evaluators should tabulate the indicator by single years or by age groups to guarantee accurate interpretation of the indicator.

Evaluators may have problems arise in measuring this indicator in settings where sexual activity outside of marriage is stigmatized, because adolescents who have initiated sex may be reluctant to admit having done so. Given the sporadic nature of sexual activity among adolescents, especially younger adolescents, in many settings, the indicator may not reflect the number/percent of adolescents who have been sexually active in the recent past (e.g., the last 3 or 6 months).

Indicator

AGE AT FIRST INTERCOURSE

Definition

This indicator is a summary measure of the average age at which adolescents become sexually active. Evaluators may alternatively calculate the indicator as: (1) the mean age at first intercourse, (2) the median age at first intercourse, or (3) the percent of youth who have ever had intercourse by selected reference ages (e.g., age 13, 15, 17, 19).

(1) Mean age at first intercourse

$$\frac{\Sigma \text{ age of adolescents}}{\text{Total \# of adolescents}}$$

(2) Median age of first intercourse[1]

$$\text{MEDIAN} = L + [(50-cf/f)*i]$$

Where:

L = the true lower limit of the class interval in which the median is located;

50 = the 50 percentile observation;

cf = the cumulated frequency up to the median class interval;

f = the frequency within the median class interval; and

i = the class width.

(3) Percent of youth who have had intercourse at reference ages

$$\frac{\text{\# of adolescents of a reference age who report having had intercourse}}{\text{Total \# of adolescents of the reference age}} \text{ x 100}$$

Data Requirements

Self-report by adolescents on whether they have ever had intercourse and, if so, their age at first intercourse. Obtaining current age is also useful for more refined measures (see below). Evaluators should measure the indicator for both married and unmarried youth.

Data Source(s)

Surveys of program participants or adolescents in the program's intended population

Purpose and Issues

The typical or average age at which adolescents in the program's intended population are initiating sex is an important parameter for program design purposes and a key outcome indicator for programs aimed at delaying onset of sexual activity. The preferred form of the indicator is the median age at first intercourse, as this form avoids bias problems that arise in the use of the mean age in settings where sexual initiation typically occurs at later ages. If fewer than 50 percent of the sample is sexually active, the preferred form of the indicator is the proportion of adolescents who had initiated sex by specified reference ages among respondents who are the reference age or older (e.g., the percentage of adolescents 16 years of age or older who had initiated sex by age 15). Evaluators may compute median ages at first pregnancy or birth in a similar fashion.

---

[1] Note: This formula is for use with grouped data consisting of percentage frequencies in each class. For ungrouped data, the median is the value of the observation falling at exactly the 50th percentile of the distribution of observations.

Gender Implications of this Indicator

Whereas menstruation is considered a sign of a young girl's passage into womanhood, in many societies, first sex marks a young man's initiation into manhood. Boys generally initiate sex earlier than girls, because many cultures tolerate or encourage sexual activity among adolescent males. In some places, a young man's masculinity is questioned if he has not had sexual intercourse by a certain age (McCauley and Salter, 1995). Responses to the question of age at first intercourse may thus be misreported because of cultural norms that may encourage boys to boast about early sexual experimentation, while having the opposite effect on girls, who may underreport sexual activity because of the great value placed on virginity.

NUMBER OF SEXUAL PARTNERS AMONG SEXUALLY ACTIVE ADOLESCENTS DURING A SPECIFIED REFERENCE PERIOD

Definition

The number of sexual partners during a specified reference period (e.g., the last 3, 6, or 12 months) among sexually active adolescents

Data Requirements

Responses to survey questions on number of sexual partners during the specified reference period

Data Source(s)

Surveys of program participants or adolescents in the program's intended population

Purpose and Issues

Having multiple sexual partners increases the risk of transmission of STIs and HIV/AIDS. This indicator provides a measure of how prevalent this high-risk behavior is in a program's intended population. However, because consistent condom use greatly reduces the risk of STI transmission, evaluators must consider the indicator in conjunction with the indicators pertaining to condom use (**Percent of Sexually Initiated Adolescents Who Used a Condom at First/Last Sex**, **Percent of Sexually Active, Unmarried Adolescents Who Consistently Use Condoms** – see below) in order to assess the prevalence of high-risk behaviors.

Because the indicator deals with a sensitive topic, there is reason to be concerned about the accuracy of reported information. Unfortunately, little methodological research has been undertaken to assess how accurately numbers of partners are reported in different settings.

NUMBER/PERCENT OF ADOLESCENTS WHO HAVE EXPERIENCED COERCIVE OR FORCED SEX

Definition

The number or percent of adolescents reporting some form of coerced or forced sex including: rape, date rape, domestic violence (resulting in sexual intercourse), sexual assault, sexual harassment, incest, and sexual molestation (Kidman, 1993)

This indicator is calculated as:

$$\frac{\text{\# of adolescents reporting forced or coerced sex}}{\text{Total \# of adolescents}} \times 100$$

Data Requirements

Self-reports of adolescents of the occurrence of coerced or forced sex either in the immediate or distant past

Data Source(s)

Surveys of adolescents in a program's intended population; program "intake" interviews; interviews during health service provision and/or in connection with educational or counseling programs

Purpose and Issues

Although most ARH programs emphasize self-efficacy and decision-making with regard to sexual relations and contraception, many adolescents, especially female adolescents, experience forced sexual encounters. At the 1994 International Conference on Population and Development in Cairo and at the 1995 Fourth World Conference on Women in Beijing, discourse on sexual and reproductive rights appropriately characterized sexual coercion as a symptom of the limited life options of girls and young women. Thus, program models designed to reduce sexual activity among adolescents must not only offer information, but must also promote public acknowledgment of the prevalence of sexual coercion and of the gender inequality that fosters it. The plausible existence of a considerable amount of coerced sexual activity highlights the inadequacy of current ARH program models, which primarily assume that sexual activity among adolescents is voluntary (Mensch, Bruce, and Greene, 1998).

This indicator provides a measure of the relative frequency of adolescents victimized by forced sex. For various reasons, incidents of coerced or forced sex are likely to be significantly under-reported in survey interviews. Evaluators can likely obtain more complete reporting in connection with counseling programs. However, in many settings, such programs reach so few adolescents, that the actual incidence is likely to be seriously under-reported. Because of the sensitivity of this matter, interviewers must often ask questions about coercive sex repeatedly to offer adolescent respondents an opportunity to disclose their experience with forced sex. One potentially effective way of broaching the subject is to ask the adolescent: "Did you have any upsetting sexual experiences in childhood or adolescence?" (Heise, Moore, and Toubia, 1995). Other researchers have also asked: "Did someone ever make you touch their breasts or genitals, or touch yours, when you did not want to?" (Boyer and Fine, 1992). After receiving a positive response, researchers or counselors can probe more deeply by asking: the age at first abuse, the frequency of occurrence, the type of abuse, whether abused by one or more people, the relationship of person(s) to the respondent, the location of the abuse, and whether the respondent told anyone else about the abuse.

Percent of Sexually Initiated Adolescents Who Used a Condom at First/Last Sex

Definition

The percent of adolescents who ever had sex reporting that they used a condom during first and/or last sexual intercourse

This indicator is calculated as:

> # of adolescents who report having
> used a condom at first/last sex ———— x 100
> Total # of adolescents

Data Requirements

Reports of condom use at first and last sexual encounters

Data Source(s)

Surveys of adolescent program participants or adolescents in the intended population for the program

Purpose and Issues

This indicator measures the prevalence of condom use at two important reference points. Reported use at first intercourse indicates the effectiveness of program mes-
sages encouraging the use of condoms among youth who become sexually active – an especially important message for programs focusing on younger adolescents. Condom use at last intercourse approximates the current condom prevalence rate among adolescents (assuming that last sexual encounters occurred in the recent past). Given that most adolescents need non-permanent methods that provide dual protection against pregnancy and STI transmission, this indicator specifies condoms. However, evaluators may include questions on the use of other contraceptive methods to capture the use of other protective methods. Ideally, evaluators will tabulate the indicator separately for married and unmarried adolescents, because the circumstances surrounding contraception and choice of method are quite different for each group. Note that reports of condom use (or the use of other contraceptives) do not necessarily reflect consistent use, measured by another indicator.

PERCENT OF SEXUALLY ACTIVE, UNMARRIED ADOLESCENTS WHO CONSISTENTLY USE CONDOMS

Definition

The percent of sexually active, unmarried adolescents who report using a condom in all sexual encounters during a defined reference period (e.g., last 6 or 12 months)

Because condom use varies by partner in some settings, the preferred approach is to ask about respondent's regular partner and recent non-regular partners (if any). The indicator measures only unmarried adolescents because of potentially confounding issues surrounding the use of condoms by married couples.

This indicator is calculated as:

$$\frac{\text{\# of unmarried adolescents who report using condoms in all sexual encounters in a reference period}}{\text{Total \# of unmarried adolescents}} \times 100$$

Data Requirements

Reports of condom use during recent sexual encounters

Data Source(s)

Surveys of program participants or youth in the geographic area of the program To measure this indicator, evaluators can use responses to questions asking whether youth "always," "most of the time," "sometimes," or "never" use condoms.

Purpose and Issues

Because condoms protect against both pregnancy and STI transmission and are readily available from non-clinic sources in most settings, many ARH programs promote condoms as the contraceptive method of choice for adolescents. However, prior research indicates that adolescents tend to be inconsistent contraceptive/condom users and/or to use condoms with non-regular partners but not necessarily with regular partners. Also, some evidence shows that the regularity of condom use tends to decline as the duration of sexual relationships increases. In view of these findings, many ARH programs counsel condom use in all sexual encounters, irrespective of the partner and duration of relationship. This indicator measures the prevalence of this "preferred" practice in a program's intended population.

PERCENT OF ADOLESCENTS WHO REGULARLY USE DRUGS/ALCOHOL

Definition

The percent of adolescents reporting that they use drugs and/or alcohol regularly

Evaluators may allow adolescents to define "regular use" or may define regular use in terms of number of times used during a specified reference period (e.g., the last month).

This indicator is calculated as:

> # of adolescents who report regular use
> of drugs and alcohol
> ———————————————————— x 100
> Total # of adolescents

Data Requirements

Responses to survey questions on drug and alcohol use
Specific questions may include:

- In the last month, how many times did you drink alcohol?
- How many times did you get drunk?
- The last time you drank alcohol, how many drinks did you have?

Data Source(s)

Surveys of program participants or youth in the program's intended population

Purpose and Issues

Studies in both the U.S. and developing countries have shown that drug and alcohol use among adolescents is associated with a higher prevalence of risky sexual behaviors (e.g., unprotected sex, multiple sexual partners). This indicator thus measures the prevalence of these non-sexual risk factors for adverse RH outcomes in the program's intended population. Where the prevalence is high, programs may need to directly address substance abuse as a proximate cause of adverse RH outcomes along with sexual and contraceptive behaviors.

PERCENT OF ADOLESCENTS WHO FEEL "CONNECTED" WITH THEIR PARENTS/FAMILY

Definition
The degree to which adolescents feel "connected" with their parents/family

Evaluators measure connections in terms of the closeness of relationships between adolescents and parents or other adult family members or caretakers.

This indicator is calculated as:

$$\frac{\text{\# of adolescents who report feeling connected to their parents/families}}{\text{Total \# of adolescents}} \times 100$$

Data Requirements
Responses to survey questions on the degree of family "connectedness" among adolescents. Evaluators may include the following types of items in indices of connectedness:

- Parents or other adult family members spend time with adolescents;
- Adolescents perceive "closeness" with parents or other adult family members;
- Adolescents perceive that they can approach parents or other adult family members with problems;
- Adolescents perceive that their parents and families care about them;
- Parents or other adult family members help with homework (for youth attending school);
- Adolescents feel comfortable talking with parents or other adult family members; and
- Adolescents perceive that they are important to their parents.

Data Source(s)
Surveys of program participants or youth in the program's intended population

Purpose and Issues
Studies in both the U.S. and developing countries have shown that feeling "connected" with parents and/or adult family members protects adolescents against risky sexual behaviors and thus against adverse RH outcomes (Resnick et al., 1997). This indicator thus measures the degree of connections between adolescents in the program's intended population and their parents/families. For diagnostic purposes, the indicator measures the percent of adolescents in the program's intended population that may be vulnerable to negative influences and to adverse outcomes. In settings where the level of connectedness is low, programs for parents and/or the provision of alternative mentors may be called for. In such programs, evaluators may also use the indicator as an intermediate outcome indicator to measure improvements in the social environment for adolescents in the program's intended population.

PERCENT OF ADOLESCENTS WHO HAVE EVER BEEN PREGNANT OR CAUSED A PREGNANCY

Definition

The percent of adolescent females who have ever been pregnant and the percent of adolescent males who have ever caused a pregnancy

This indicator is calculated as:

$$\frac{\text{\# of adolescent females who report having ever been pregnant}}{\text{Total \# of adolescent females}} \times 100$$

$$\frac{\text{\# of adolescent males who report having ever caused a pregnancy}}{\text{Total \# of adolescent males}} \times 100$$

Data Requirements

Responses to survey questions asking whether female adolescents have ever been pregnant and whether male adolescents have ever caused someone to be pregnant

Data Source(s)

Surveys of program participants or youth in the program's intended population

Purpose and Issues

Many ARH programs aim to reduce the number of adolescent pregnancies. This indicator provides a simple proxy measure of the level or volume of adolescent pregnancies for use in assessing the impact of such programs.

Alternatively, evaluators can use a pregnancy rate (i.e., number of annual pregnancies per 1,000 women 15-19 years of age). Evaluators can calculate this alternative measure in the same manner as they calculate the age-specific fertility rate for adolescents (see Part III.B), but they will base the calculation upon pregnancies instead of on live births.

In settings where sexual relations and pregnancy outside of marriage are highly stigmatized, female respondents to surveys will likely under-report adolescent pregnancies occurring outside of marriage. Responses by male adolescents may also be biased, but the direction of the bias is less certain. On the one hand, males who have had multiple casual sexual partners may be unaware of pregnancies they caused. On the other hand, male adolescents may exaggerate their sexual prowess in surveys, and thus may over-report the number of pregnancies they caused. Nevertheless, in most settings, the indicator provides a lower-bound estimate of the true percent of adolescents experiencing or causing pregnancies.

Although ARH programs tend to view adolescent pregnancies as a negative outcome, pregnancies occurring to adolescents are sometimes wanted. Some situations may provide economic and social benefits of pregnancy during adolescence. Thus, evaluators must interpret this indicator in conjunction with data on the "wanted" status of pregnancies occurring to adolescents.

Unmet Need for Family Planning Among Adolescents

Definition
The number/percent of sexually active adolescents who do not desire to become pregnant immediately but who are not using a contraceptive method

Data Requirements
Responses to survey questions on:

- Desire for (additional) children, and if so, the desired time until pregnancy/birth;
- Current contraceptive use status;
- Current sexual activity, fecundity, pregnancy, and amenorrhea status for women not currently using a contraceptive method;
- Wanted status of last pregnancy for currently pregnant or amenorrheic women; and
- Use of a contraceptive method at the time of the current/last pregnancy.

Data Source(s)
Surveys of adolescents in the program's intended population

Purpose and Issues
Although many programs operate on the assumption that adolescents desire to avoid pregnancy, becoming pregnant is advantageous in some settings and situations. The present indicator provides a measure of unmet need for family planning comparable to the generic indicator (Bongaarts, 1990; Westoff, 1991) but adjusted for the situation of many adolescents. The adjustments include sexually active female adolescents who are not currently married or in union. Evaluators may also survey for "sexually active" adolescents in terms of all women reporting ever having had sexual intercourse. However, this latter approach will include women who have sex only sporadically; the fact that they are not currently using a contraceptive method may overstate the level of unmet need for family planning. Alternatively, evaluators may use reported contraceptive use at last sexual intercourse in lieu of current contraceptive use.

Percent of Adolescents Who Were Ever Diagnosed with an STI

Definition
The percent of adolescents who have ever been diagnosed as having an STI

This indicator is calculated as:

$$\frac{\text{\# of adolescents ever diagnosed with an STI}}{\text{Total \# of adolescents}} \times 100$$

Data Requirements
Responses to survey questions on whether adolescents had ever been diagnosed as having an STI Questions include:

- Has a physician or nurse ever told you that you had an STI?; and
- Have you ever tested positive for an STI?

Data Source(s)
Surveys of program participants or youth in the program's intended population

Purpose and Issues
Along with reducing the incidence of adolescent pregnancy, reducing the incidence of STIs among adolescents is an important objective of many ARH programs. This indicator provides a relevant long-term outcome measure for such programs. Because the indicator is a "lifetime" measure, it does not measure incidence or prevalence for specific reference periods, although evaluators can derive an incidence-like measure by obtaining information on the dates of episodes of STIs. The measure is, however, crude and suffers from several biases. First, many STIs lack recognizable symptoms; thus, the indicator will underestimate the true percent of adolescents who have ever contracted an STI. Second, because STIs imply that respondents have engaged in behaviors stigmatized in many settings, the indicator is prone to under-reporting in survey interview situations. Thus, the indicator will provide a lower-bound estimate in most settings.

**Part III.I
Postabortion
Care**

- Legal status of abortion

- Policy status of abortion

- Abortion rate (AR) and total abortion rate (TAR)

- Total number of hospital admissions for abortion-related complications

- Number/percent of SDPs providing postabortion care services by type and geographic distribution

- Number/percent of practitioners trained in PAC by type and geographic distribution

- Percent of SDPs providing postabortion care services that meet a defined standard of quality

- Number/percent of SDPs that offer family planning to postabortion care patients

Complications of unsafe abortion are a major contributor to maternal mortality and morbidity in developing countries and have been recognized by the international community as a key public health issue. In recent years, national Ministries of Health, NGOs, international reproductive health agencies, and donor organizations have increased their efforts to improve access to high-quality postabortion care (PAC). PAC generally includes clinical treatment for complications of incomplete abortion, provision of counseling and contraceptive supplies (to avoid a repeat abortion), and in some locations, referral to other reproductive health care. PAC may also include community education to improve reproductive health and to reduce unwanted pregnancy and the need for abortion.

All but a few countries legally permit induced abortion under one or more circumstances: to save a woman's life; in cases of rape or incest; to protect a woman's mental or physical health; or to meet socioeconomic needs. Most donor-supported programs, including those assisted by the U.S. Agency for International Development (USAID), prohibit the promotion of abortion as a method of family planning. This *Compendium* focuses on use of the indicators for monitoring and evaluating postabortion care.

For most indicators in other sections of this *Compendium*, the "desired direction" of the indicator is clear. For example, in HIV prevention, one seeks increased use of condoms and decreased incidence of HIV infections. By contrast, increased use of postabortion care services is a more ambiguous indicator of the effectiveness of these services. A good postabortion care program may treat an **increasing** number of cases in the short term; this increase may indicate that improved services are leading more women with complications to avail themselves of the services, or it may mean that the number of poorly performed abortions has increased in the community. (One expects that postabortion or induced abortion care will decline in the longer term as women begin to use contraception more effectively). In light of these considerations, the best indicators for effective postabortion care are increases in the availability and quality of services rather than increases in the use of these services.

As a final note, one indicator in this part of the *Compendium*, the rate of abortions, could serve purposes other than evaluating postabortion care. A key objective of many family planning and reproductive health programs is to reduce the rate of abortions through provision of family planning services to prevent unwanted pregnancies (Westoff and Bankole, 2001). Reductions in the abortion rate may signal improvements in access and quality of care in family planning services, although reductions in abortion may lag behind if women are motivated to limit fertility but are not yet using contraception.

The reproductive health community has less experience using and testing the indicators in this section than it does those in most other sections of the *Compendium.* However, as new PAC initiatives emerge, the need for accountability is as great in this area as in other areas of reproductive health. Thus, we present a number of indicators in the spirit of stimulating dialogue and identifying those most useful to program managers working in this area.

Methodological Challenges of Evaluating PAC

Ideally, evaluators would like to monitor postabortion care using such indicators as availability and quality of care, the rate and number of cases treated for abortion-related complications, adoption of family planning by postabortion patients, numbers of unsafe abortions and repeat abortions, and the number and type of referrals to other reproductive health services. Ultimately, evaluators and researchers would also want to assess the impact of improved services on maternal mortality and morbidity. Such information is virtually nonexistent in the vast majority of developing countries. Most countries have little evaluation infrastructure and face pressing service delivery needs that often take priority over monitoring, evaluation, and accurate reporting.

- **Data on trends and consequences of abortion and on the prevalence of unsafe abortion are generally difficult to collect.**

Some national-level reproductive health surveys (e.g., DHS and RHS) have asked questions about abortion, but the data have proven unreliable in most cases. Information on the risks of unsafe abortion for some vulnerable populations is particularly elusive. Certain women at high risk for unsafe abortion, such as adolescents and women who are refugees/internally displaced, may not seek services in the public sector or may have limited access to such services. As such, the available data do not cover these groups. In particular, it would be useful to know more about contraceptive use and the magnitude and consequences of unsafe abortion among these groups.

- **Where existing systems for monitoring reproductive health programs exist, they often exclude items related to postabortion care.**

For example, in many countries, the commodities/logistics system or medical supplies list covers commodities related to all aspects of reproductive health programs; yet, rarely do such systems track the procurement and use of manual vacuum aspiration (MVA) instruments used in treating abortion complications.

- **Even a simple count of the number of public and private facilities that provide PAC is often difficult to obtain.**

To avoid unwanted attention from higher authorities or to guard the anonymity of the patients involved, many facilities systematically avoid reporting the number of cases treated for the complications of abortion under the mistaken belief that PAC is illegal. To further complicate the problem, abortion-related cases are often classified as hemorrhage or infection, and are thus difficult to identify as abortion-related.

- **The private sector in many countries treats the complications of abortion; yet, little systematic reporting is available.**

In fact, service providers may not openly "advertise" that they provide these services. This scarcity of reliable data from both public and private sector services greatly impedes the monitoring and evaluation process; evaluators will need to make concerted efforts in most countries to improve data collection.

- **Information on the quality of abortion-related care is also difficult to collect.**

Ideally, evaluators would like to collect data on the technique used for uterine evacuation, length of patient stay, adequacy of postabortion contraceptive counseling, provision of a method, patient satisfaction, among others. In fact, methodologies and data collection instruments exist for assessing quality of care (once one identifies facilities providing the services and has permission to conduct an evaluation of the services). Yet even then, the techniques available for assessing quality for more routine reproductive health services (such as the Service Provision Assessment [SPA] or the Quick Investigation of Quality [QIQ] for family planning, described in Part II.H.2) may be more difficult to apply in the case of postabortion care, given the desire for patient confidentiality and the psychological distress a woman may be experiencing. In addition, assessment of services is made more difficult by the round-the-clock, emergency nature of the services and the relative low frequency of PAC patient arrivals compared to family planning visits or deliveries in maternity settings. Furthermore, most developing country health systems lack guidelines and protocols for postabortion care, and the international guidance that exists currently is limited and poorly disseminated (WHO, 1995b; Rogo, Lema, and Rae, 1999).

- **Determining the impact of unsafe abortion on maternal morbidity and mortality is rarely feasible.**

Reproductive health programs seek to save the lives and protect the health of women who undergo unsafe abortion. Yet tracking the impact of unsafe abortion on maternal mortality and morbidity is fraught with a double set of challenges. The first relates to trying to measure abortion accurately. The second is measuring maternal mortality and morbidity with precision and assigning causes, including unsafe abortion. (For a more complete description of the methodological problems relative to the measurement of maternal mortality, see Part III.D). In short, it is difficult to measure maternal mortality and morbidity, much less the contribution of unsafe abortion to these two outcomes, or of interventions to reduce unsafe abortion. The WHO is the major agency that has provided estimates on maternal mortality at regional and global levels and on the contribution of unsafe abortion, although the unsafe abortion estimates have not been updated recently (WHO, 1998c).

Programs can address some of the challenges to monitoring and evaluation in this area in the short term through dedicating greater financial and human resources to collecting and analyzing data. The indicators discussed below are designed to provide a starting point for these efforts. Addressing other challenges identified above will depend on developments in the larger environment of policies, medical practices, and cultural attitudes that determine how health systems respond to unsafe abortion.

LEGAL STATUS OF ABORTION

Definition

The legal restrictions that establish the circumstances under which a woman can legally terminate a pregnancy

Five possible degrees of restrictiveness of abortion laws can exist for a given country (CRLP, 1999):

1) Abortion is permitted without restriction as to reason;
2) Abortion is permitted on socioeconomic grounds (such laws allow consideration of a woman's economic resources, her age, her marital status, and the number of her living children);
3) Abortion is permitted to protect a woman's mental health, as well as her life and physical health (interpretation of "mental health" may vary across countries, but it may encompass, for example, the psychological distress suffered by a woman who is raped or the severe strain caused by socioeconomic circumstances);
4) Abortion is permitted to protect a woman's life and physical health (such laws may permit abortion on health grounds that may require the threatened injury to health be either serious or permanent); and
5) Abortion is permitted only to save a woman's life, or the procedure is banned entirely.

Data Requirements

Text of existing laws

Data Source(s)

Penal codes, health codes, as well as reports of international law organizations that monitor the status of abortion (e.g., Center for Reproductive Law and Policy website, www.crlp.org)

Purpose and Issues

The purpose of this indicator is to measure the degree to which a woman has access to safe abortion care and postabortion care in a given country. However, abortion laws are just one factor, albeit an important one, which influence access to care. Various policies and the manner in which they are implemented, for example, may be more critical factors than the law(s). (See the following indicator, **Policy Status of Abortion**.) Furthermore, although almost all countries have at least one legal indication (reason) for abortion, induced abortion services may be unavailable to the extent allowed by law.

Gender Implications of this Indicator

An estimated 20 million unsafe abortions occur each year, and unsafe abortion contributes an estimated 13 percent of total maternal mortality, or 76,000 deaths per year (Ipas, 2000). The great majority of deaths from abortion occur in countries where abortion is either illegal, or where abortion is legal, but the status and access uncertain enough (such as in India) that women still resort to unsafe abortion. Maternal mortality has been identified as a compelling gender equity and human rights issue, and reduction of maternal mortality is called for in the action plans of numerous international conferences and conventions. It is difficult to reduce maternal mortality without attention to the toll that unsafe abortion takes. Where abortion is illegal, access to high-quality postabortion care, including provision of family planning, is critical.

Indicator

POLICY STATUS OF ABORTION

Definition

The policy environment concerning abortion

Policies are defined as including abortion laws but also regulations, guidelines, financial provisions, and customary practices affecting the delivery of abortion and postabortion care. These policies may be written or unwritten (e.g., informal guidelines) and may clarify how program services should be operationalized and provided at the health system level.

Evaluators can rate the policies on abortion care in terms of four levels of restrictiveness:

1) Policies that encourage wide availability of abortion care, with few restrictions on access;
2) Policies that allow abortion care to be provided, but with some restrictive provisions;
3) Policies that significantly restrict access to abortion care; and
4) Policies that prohibit provision of abortion care, except in relatively rare situations.

Similarly, evaluators can classify the level of support for postabortion care by four levels:

1) Policies that are highly favorable towards the treatment of complications from abortion, including complications from illegal abortion;
2) Policies that are moderately favorable towards postabortion care;
3) Policies concerning postabortion care that are virtually non-existent, and care provided is *ad hoc* and highly variable; and
4) Policies that limit women's access to postabortion care.

Data Requirements

Information on current policies

Data Source(s)

Penal codes; special statutes; court decisions; public health regulations and administrative codes; medical and nursing standards; and health care facility protocols

Note: While some policies are written and widely disseminated, others are informally developed and implemented by health facilities and providers as a matter of customary practice. For these informal policies, interviews with health care providers and administrators may enable one to determine the prevailing policies in a given geographical setting.

Purposes and Issues

Laws specifically addressing when a pregnant woman can have an abortion based on the circumstances of her pregnancy provide an initial indication of the policy environment. Currently, about 62 percent of the world's population live in the 64 countries that legally permit abortion either without restriction as to reason or on broad socioeconomic grounds. The remaining 38 percent live in countries that have varying degrees of restriction (CRLP, 1999).

Yet abortion and postabortion care-related laws can be highly misleading with regard to what authorities actually tolerate or encourage, which may be either more or less restrictive than the laws alone would indicate. In some countries, "menstrual regulation," or early abortion when pregnancy has not been confirmed, is officially sanctioned and widely available, even though laws on the books are highly restrictive with respect to "abortion." In other countries, abortion is legal for a wide range of circumstances, but policies significantly restrict access. For this reason, one must often go beyond the official regulations and determine the actual policy toward abortion.

A range of institutions and individuals (e.g., clinical providers, administrative health care personnel, professional associations, judicial authorities) may create and implement policies, which may vary by geographic area or from one health care system/facility to another. Furthermore, policies with a major impact on postabortion care and abortion care may not be formally spelled out, but rather may be developed and carried out on an *ad hoc* basis.

Examples of the content included in abortion-related policies include:

- Circumstances under which a pregnant woman is permitted to have an abortion, including her age, marital status, circumstances of the pregnancy, and length of the pregnancy;

- Types of health care facilities authorized to provide induced abortion or postabortion care;
- Types of health care providers authorized to provide induced abortion or postabortion care;
- Funding of abortion-related services (e.g., availability of publicly funded subsidies for services, availability of sliding-scale fees, requirements for patient purchase of medications and supplies);
- Waiting periods, provider signatures, spousal or parental consent, and other procedural requirements prior to service delivery; and
- Types of clinical instruments and medications approved.

ABORTION RATE (AR) AND TOTAL ABORTION RATE (TAR)

Definition
The abortion rate is the number of induced abortions occurring in a specified reference period (e.g., one year) per 1,000 women of reproductive age (15-44 or 15-49).

The abortion rate (AR) is calculated as:

$$\frac{\text{\# of abortions}}{\text{Total mid-year population of women}} \times 1000$$
$$\text{15-44 (or 49)}$$

The total abortion (TAR) rate is the total number of abortions a woman will have in her lifetime if current levels persist. This lifetime risk is a cohort measure and can be calculated with period measures (age-specific abortion rates) or approximated by multiplying the abortion rate by the length of the reproductive period (30-35 years), (Bertrand and Tsui, 1995). Thus, the total abortion rate is calculated as:

$$\text{TAR} = 35 \times \text{abortion rate}$$

Where:
35 = # of years of reproductive life span.

Data Requirements
Total number of induced abortions occurring in a given year or reference period; the enumerated or estimated mid-period population for the same period

Data Source(s)
Data on abortions: where abortion laws are liberal, official statistics are likely to provide the most accurate numbers; where abortion is restricted, data will be less accurate but one may derive estimates from surveys of providers, population-based surveys, hospital-based studies, or a combination of sources.

Census data or projections based on census data usually provide information on the population of women 15-49.

Purpose and Issues
Rates and ratios are two of the most widely used abortion measures. They are indispensable statistics for documenting levels of abortion across time and space. The abortion rate is a useful tool in the evaluation of contraceptive services, either for the purpose of setting a baseline or for measuring progress. The abortion rate reflects contraceptive method and user effectiveness, as well as access to services. The rate is less useful than facility-based or other data for the evaluation of demonstration projects or the effects of separate program components.

Several factors affect the rate: 1) the proportion of women who become pregnant in a year; 2) the likelihood a pregnancy is unwanted; and 3) the likelihood an unwanted pregnancy will be terminated. Consequently, increasing effective contraceptive use and thus decreasing the number of unwanted pregnancies can lower the abortion rate, while potentially lowering fertility. On the other hand, if the number of pregnancies is constant but more unwanted pregnancies are carried to term, the abortion rate will also decrease, with the effect of potentially increasing fertility.

Like the total fertility rate, the total abortion rate is easily understood and serves as an effective statistic for comparative purposes. The advantage of the TAR is that is takes into account the probability of becoming pregnant and the probability of terminating each pregnancy throughout the reproductive life cycle. Like the abortion rate, a high TAR may indicate several factors, including the availability and quality (or lack thereof) of contraceptive services. A high TAR may also reflect a high prevalence of traditional contraceptive method use in a given country.

Development of the abortion rate by specific age ranges (such as 15-19 or 24) can also be a useful tool for documentation of those groups at particular risk for abortion.

Rates and ratios, however, are often seriously compromised in terms of accuracy. As stated above, where abortion is restricted, data are likely to be inaccurate. In these circumstances, data may require adjusting for underreporting, misclassification or socioeconomic conditions that reflect the safety of clandestine abortion and the likelihood that a woman with complications from an induced abortion seeks and receives treatment. Even in less restricted settings, research has shown that women underreport their abortion experiences. Where judicial reprisal or severe physical or psychological damage is a possibility, accurate reporting is even less likely. For years, researchers have attempted different methods for collecting sensitive and personal information at little or no threat to the research subjects. Recent efforts with the National Survey of Family Growth in the United States include computer-assisted interviews, which maximize a respondent's privacy.

TOTAL NUMBER OF HOSPITAL ADMISSIONS FOR ABORTION-RELATED COMPLICATIONS

Definition

The total number of admissions to a health care facility for abortion-related complications for a reference period (e.g., one-year)

This indicator includes both complications resulting from spontaneous abortion (miscarriage) and those occurring as a result of induced abortions.

Postabortion complications include hemorrhage, local and systemic infection, injury to the genital tract and internal organs, and toxic or chemical reactions from attempts at self-induced or unsafe abortion. This indicator omits long-term sequelae (physical impairment, pain, pelvic inflammatory disease, secondary infertility, increased rate of ectopic pregnancy).

Data Requirements

Counts of women admitted to a health care facility for treatment of abortion-related complications during a reference period

Data Source(s)

Special studies or services statistics from health facilities providing treatment of abortion complications

Note: In hospitals in developing countries, treatment of abortion complications may be performed in many different locations within the facility, such as the gynecological ward, emergency room or operating room; data collection should therefore include admissions from all locations.

Purpose and Issues

This indicator monitors changes in caseloads and has important administrative implications. Evaluators and managers can also use it to track resource use and needs for treatment of abortion-related complications. It also has policy implications in that it is useful for assessing the cost of unsafe, induced abortions to individual hospitals or to a national health system. Numbers of admissions for abortion complications can also provide

denominators for other useful indicators, such as the percentage of PAC patients under the age of 20 or the percentage of PAC patients presenting at 12 or fewer weeks of pregnancy. In some individual facilities, such as health centers, however, the number of admissions for abortion complications may be small so that calculation of percentages may be inappropriate.

Peru provides an example of current efforts to improve the quality of information about PAC caseloads at public sector health centers and hospitals in one state. The process involves completion of a standardized clinical history form for each postabortion patient receiving treatment in the facility. The form requests information about a limited number of key indicators, such as diagnosis, patient age, evacuation technique used, and duration of pregnancy. Providers in these facilities are accustomed to completing a similar form for obstetric deliveries, so they have easily adopted the form. Staff are responsible for entering the information into a database at each facility, and the Ministry of Health makes the information available for use at the facility, state, and national levels. The Ministry of Health and Ipas are assessing this pilot project for possible use in other states in Peru and, eventually, in other countries.

This indicator includes both complications due to induced and to spontaneous abortions. While it is often of interest to distinguish between the two types in order to estimate the number of induced abortions, this information is often difficult to obtain. Moreover, many would question the ethics of asking young women if they have had an abortion in restrictive legal settings. Clinical evidence is often inconclusive, and reports may also be heavily biased in restrictive environments. Even where service providers are fairly certain that an abortion-related complication results from an induced abortion, they may choose not to report this in the records due to a legally and/or socially restrictive environment. This omission results in service data that are potentially misleading in terms of the number of spontaneous versus induced abortions.

Evaluators have used several approaches in attempting to distinguish between spontaneous and induced abortions. These approaches range from a series of questions asked of the patient to multipliers based on the biological occurrence of spontaneous abortion to multipliers based on expert opinion of the proportion of hospitalizations due to complications of induced abortion.

This indicator can estimate the extent of induced abortion in countries where abortion is restricted. Researchers have utilized data on abortion-related hospital admissions to construct such estimates. Evaluators can extrapolate the number of abortion-related hospital admissions to estimate the number of abortions in the population by using a variety of multipliers. These multipliers will be region and country specific. They will vary by the degree of restrictiveness of the legal and social climate, the availability of induced abortion performed by trained providers, the procedures used by clandestine providers, the availability of antibiotics, and the socioeconomic status of the women who undergo abortions. To address the uncertainty related to these multipliers, researchers have suggested using a range of estimates with several different multipliers (Singh and Wulf, 1994).

The best way to collect data for this indicator may be to conduct special studies at specific facilities (e.g., hospitals in urban areas). A hospital-based study in Nigeria indicated that over 75 percent of gynecological admissions to hospitals were due to abortion-related causes (Rogo, Lema, and Rae, 1999).

Possible alternative indicators include:

- Percent of cases of obstetric complications treated at service facilities that are abortion-related; and
- Percent of women treated for PAC at service facilities who die.

NUMBER/PERCENT OF SDPs PROVIDING POSTABORTION CARE SERVICES BY TYPE AND GEOGRAPHIC DISTRIBUTION

Definition

The total number and percent of service delivery points (SDPs) providing PAC services by type of facility (e.g., health center, district hospital, private physician) and geographic location

Service delivery points should include those in both the public and private health care sectors.

This indicator is calculated as:

$$\frac{\text{\# of SDPs of a particular type that deliver PAC services in a given area}}{\text{Total \# of SDPs of that type in the area}} \times 100$$

Postabortion care consists of:

- Emergency treatment of complications from spontaneous or unsafely induced abortions;
- Family planning counseling and services; and
- Linkages to comprehensive reproductive health care.

The need for community education to reduce the need for abortion and improve reproductive health is often considered a fourth element of postabortion care.

Data Requirements

Total number, type, and geographic location of facilities providing postabortion care services; total number of service delivery facilities by type and location

Data Source(s)

National program records; private and NGO records; provider interviews; and observation of services

Purpose and Issues

The purpose of this indicator is to measure the degree to which postabortion care services are available within a given country. All countries should be able to monitor the availability of PAC services for treatment of abortion complications. In countries where abortion is severely restricted, evaluators may have difficulty obtaining accurate information on all the facilities providing postabortion care services. Even in countries where abortion laws may be less restrictive, providers (especially private providers) may be less open to admitting that they provide such services because of the stigma attached to abortion and a desire to protect the privacy of their patients.

Ideally, information collected for this indicator can serve to monitor other key indicators, such as the percentage of facilities offering PAC services in a given region or at a given level of care (e.g., health center, district hospitals, and tertiary hospitals). Another alternative indicator is the number, type, and geographic distribution of SDPs that have commodities, equipment, and transport for postabortion care.

If population figures are available, information collected for this indicator may help determine if the number and type of facilities providing services are sufficient for the population served. Indicators developed by UNICEF, WHO, and UNFPA for monitoring access to emergency obstetric care can provide guidance.

NUMBER/PERCENT OF PRACTITIONERS TRAINED IN PAC BY TYPE AND GEOGRAPHIC DISTRIBUTION

Definition
The total number and percent of practitioners trained in PAC by type of specialty and geographic location

This indicator is calculated as:

$$\frac{\text{\# of practitioners of a particular specialty and practicing in a given area trained in PAC}}{\text{Total \# of providers of that specialty practicing in the area (e.g., district)}} \times 100$$

Postabortion care consists of:

- Emergency treatment of complications from spontaneous or unsafely induced abortion;
- Family planning counseling and services; and
- Linkages to comprehensive reproductive health care.

Data Requirements
Number of practitioners trained in PAC by specialty and geographic location; the number of all providers by specialty and location of practice

Data Source(s)
National program records; private and NGO program records; and training institution records (e.g., medical and midwifery school records)

Purpose and Issues
The purpose of this indicator is to measure the extent to which practitioners are trained in PAC, which in turn influences the availability of such services.

If denominators by type of practitioner are available, data collected for this indicator may be used to determine the percentage of provider types trained in a given country, for example, the percentage of ob-gyn or general practitioners or nurse-midwives trained in PAC.

A limitation of this indicator relates to the follow-up monitoring of practitioners applying their skills to their jobs. Training of providers often takes place in settings such as teaching hospitals. Once providers are assigned to their posts, they may find it challenging to apply the skills they learned. Once providers are spread out geographically in a given country, evaluators may have difficulty routinely monitoring how and whether they are using the skills they learned in training.

Alternative indicators may include:

- Number and percent of medical and nursing schools incorporating PAC into their curricula; and
- Number and percent of practitioners trained in PAC during pre-service and in-service training.

PERCENT OF SDPs PROVIDING POSTABORTION CARE SERVICES THAT MEET A DEFINED STANDARD OF QUALITY

Definition
A framework of fundamental elements necessary for the delivery of high-quality postabortion care appears in Table III.I.1 (Leonard and Winkler, 1991 and Ipas/PRIME, 1998). Developed by Ipas, this framework provides guidance on the development and use of indicators to monitor and evaluate the quality of postabortion care. Following is a list of the key components of postabortion care.

- Appropriate technologies for uterine evacuation;
- Technical performance in PAC services;
- Patient-provider interaction;
- Information and counseling on PAC; and
- Availability of equipment, supplies, and medications.

This indicator is calculated as:

$$\frac{\text{\# of SDPs providing PAC services that meet a defined quality standard}}{\text{Total \# of SDPs offering PAC services}} \times 100$$

Data Requirements
Scores for each facility on data collection instruments developed to assess the quality of services at a network of facilities; for the denominator, counts of SDPs offering PAC services.

Data Source(s)
Checklists; observation of services; informal patient and provider interviews; and review of facility records such as surgical logbooks for routine monitoring or supervision of the quality of PAC services

Facility audits, patient interviews, and/or observation of patient-provider interaction for more in-depth special studies

Purpose and Issues
Although measurement of service quality is challenging, managers and providers can use the indicators outlined in this *Compendium* and other existing resources to routinely assess their own programs. The type of services available will depend on the level of the facility within the broader health care system. Health care systems should strive to decentralize PAC services so that women have access to emergency care within a reasonable time and distance from their homes.

Researchers have conducted a number of PAC operations research projects in Latin America and Africa, taking care to ensure informed consent and protection of patient confidentiality in studies on abortion-related services. This research has documented changes in quality after the introduction of PAC interventions in a variety of health care settings.

The complexity of clinical service provision requires the monitoring of many aspects of care. Appropriate infection prevention practices, for example, require that providers carry out a number of different steps in the process of clinical care. Table III.I.1 includes possible indicators for measuring each of the key quality components of PAC. These indicators vary in their format and will require further testing to determine the most useful and valid means of assessing service quality. Service delivery staff and managers should determine how they plan to measure their program's performance. Two references provide guidance on using checklists for many of the indicators listed in Table III.I.1. (Otsea et al., 1999; Population Council, 2000).

Table III.I.1 Possible Indicators by Key Component of Quality for PAC

Indicator	Facility Audit	Observation	Provider Interview	Patient Exit Interview	Service Statistics
Appropriate Technologies					
• Availability of safe, effective technologies for uterine evacuation	✓	✓			
Technical Performance					
• Availability of providers trained in PAC	✓	✓	✓		
• Adherence to current/updated practices for uterine evacuation		✓	✓		
• Percent of appropriate patients (12 or less weeks of pregnancy) treated with MVA		✓			✓
• Adherence to current/updated infection prevention practices (universal precautions)		✓	✓		
• Adherence to current/updated pain management practices		✓	✓		
• Existence of hygienic service delivery environment	✓	✓			
• Average length of time from patient arrival at the health care facility until treatment		✓		✓	
• Mechanism for stabilization, referral, and transport of patients with severe abortion complications	✓		✓		
Patient-Provider Interaction					
• Maintenance of confidentiality of patient records and information exchanged between providers and patients		✓	✓		
• Respect for women's needs for privacy during treatment	✓	✓		✓	
• Adherence to practices of informed choice for patient decisions		✓		✓	
Information and Counseling					
• Percent of patients who receive information about their diagnosis, treatment, prognosis, follow-up care necessary, and post-treatment complications		✓		✓	

Indicator	Facility Audit	Observation	Provider Interview	Patient Exit Interview	Service Statistics
• Percent of PAC patients who receive FP counseling that includes at least the following information: ▪ Pregnancy can occur immediately; ▪ Use of a contraceptive method can delay a subsequent pregnancy if the woman wishes a method; and ▪ Location where she can obtain a contraceptive method.				✓	
• Percent of PAC patients who wish to receive a contraceptive method at the time of treatment and who receive the following information: ▪ Assessment of the woman's personal situation; ▪ Contraceptive options; ▪ Method use; ▪ Side effects of the method selected; and ▪ Resupply options.		✓		✓	
• Percent of PAC patients who receive contraceptive method prior to departure from the health care facility		✓		✓	✓
• Percent of PAC patients at risk of HIV/STI referred for testing and counseling		✓		✓	
Equipment, Supplies and Medications					
• Availability of supplies, equipment, and medications for uterine evacuation, including materials for infection prevention and pain management	✓		✓		
• Availability of a range of contraceptive methods	✓		✓		

NUMBER/PERCENT OF SDPS THAT OFFER FAMILY PLANNING TO POSTABORTION CARE PATIENTS

Definition

The total number and percent of SDPs offering family planning counseling and methods to postabortion care patients

The numerator includes all SDPs offering family planning after a woman receives treatment of abortion complications. The denominator includes the total number of SDPs offering PAC in a given country.

This indicator is calculated as:

$$\frac{\text{\# of SDPs offering family planning counseling and methods to postabortion care patients}}{\text{Total \# of SDPs offering PAC}} \times 100$$

Data Requirements

Count of the total number of centers/hospitals offering PAC and total number of centers/hospitals routinely offering family planning to women who have received PAC services

Data Source(s)

Service statistics; provider and patient interviews; and observation of services; interviews with patients or actual observation of services (preferred); provider interviews (useful); logbooks and patient records (potentially useful but often incomplete or inaccurate)

Purpose and Issues

The recovery period after PAC services is an opportunity for health providers to offer comprehensive reproductive care.

Providers should offer counseling on family planning following any postabortion care services, because women can regain their fertility as soon as 14 days after an abortion and before the next menses. Every woman treated for complications from an abortion should be aware of the following before leaving the health facility: pregnancy can occur immediately and use of a contraceptive method can delay a subsequent pregnancy if the woman wishes a method. FP methods should be available on site; and for those women who select a method at the time of treatment, counseling should include: assessment of the woman's personal situation, contraceptive options (almost all methods can be started immediately following PAC procedures), method use, side effects of the method selected, and resupply options. In countries where STI/HIV/AIDS is prevalent, providers should encourage dual protection (Rogo, Lema, and Rae, 1999).

Possible alternative indicators may include:

- Percent of PAC patients who receive family planning counseling at the time of service;
- Percent of PAC patients who accept a contraceptive method at the time of service; and
- Percent of PAC patients at risk of HIV/STI referred for testing and counseling

Part III.J
Male Involvement
in Reproductive
Health Programs

- Number of visits to male-focused services, by type of service
- Percent of men (husbands) who are supportive of their partner's reproductive health practices
- Percent of contraceptive method use requiring male cooperation
- Percent of men and women who discuss reproductive health issues with their spouse or sexual partner

MALE INVOLVEMENT IN RH PROGRAMS

Reproductive health programs have traditionally focused on women. However, since the ICPD Conference of 1994, programs have paid increasing attention to male involvement in reproductive health services. The Cairo Plan of Action highlighted the need to "promote gender equality in all spheres of life… and to encourage and enable men to take responsibility for their sexual and reproductive behavior and their social and family roles" (United Nations, 1994). In recent years, family planning and other reproductive health programs have increased their efforts to involve men in their programs, because they recognize that men have an important influence on women's and children's health, as well as distinct health needs of their own. Many advocates for involving men in reproductive health programs believe that involving men can lead to greater gender equity.

Gender is often misunderstood as a synonym for women's issues, but gender refers to the socio-cultural roles of both women and men. Men as well as women face gender-related barriers to reproductive health. Gender-sensitive programs recognize that gender inequalities between women and men significantly influence the sexual health of both men and women. Programs that aim to improve the status of women must also recognize that men influence the status of women's health.

Gender norms that act as barriers to women's involvement in reproductive health programs also act as barriers for men. Men are frequently described as the forgotten clients, particularly for family planning services. Men face both socio-cultural barriers and institutional barriers to involvement.

Socio-cultural definitions of masculinity may make it difficult for men to seek reproductive health information or services. Men who wish to limit their family size often face gender norms that equate number of children with virility and that discourage men from using reproductive health services of any kind. Some societies encourage men to take sexual risks, such as frequenting sex workers. Some have suggested that risk-taking behavior also extends to having sex without condoms (Foreman, 1999). Some cultures equate masculinity with exercising power over women. Men's fears of losing power (often triggered by women becoming primary decision-makers about family planning) can lead to gender-based violence.[1]

Men who do wish to use available services also encounter gender barriers. Men's involvement in the health system often stops at the door to the clinic. When they accompany their partners to a facility, men may find an absence of programs that encourage or allow them to participate. Men are concerned about preventing and treating sexually transmitted infections (STIs), yet often they do not know where treatment is available. Men also worry about impotency and infertility; they are more likely than women to work in areas with environmental or occupational hazards that can lead to reproductive health problems. However, clinics and programs, especially in remote and rural areas, focus on women's health.

Approaches and Objectives of Male Involvement Initiatives

Greene (1999) describes the evolution in initiatives aimed at increasing male involvement in reproductive health – specifically, she outlines four approaches to male involvement:

The traditional family planning approach: This approach, which dominated the field prior to Cairo ICPD in 1994, predominantly focuses on providing contraceptive methods to women for the purpose of reducing fertility.

The men and family planning framework: This approach assumes that men can prevent women's contraceptive use and that they themselves are an untapped group of potential contraceptors. It treats males primarily as a means of increasing contraceptive prevalence.

---

[1] See Part III.K for more discussion on violence.

The male equality framework: This approach is characterized by programs designed to serve men as RH clients in much the same fashion as programs have served women, consistent with the Cairo Programme of Action that called for increased attention to the individual needs of women, men, and adolescents.

The gender equity framework: This approach addresses the relationship between women and men and the sharing of responsibility and action. It focuses on men as supportive partners of women, and thus reflects the spirit of the ICPD document and the transformation of social roles that constrain reproductive health and rights. It emphasizes the ways services are provided and the opportunities to deliver and reinforce messages supportive of gender equity rather than specifying which RH services should be provided and to whom.

RH initiatives directed toward women generally have clearly defined behavioral objectives: to increase contraceptive prevalence, to increase the use of a skilled birth attendant at delivery, to increase the use of condoms for STI/HIV prevention, to increase the prevalence of breastfeeding, and so forth. In contrast, no single common behavioral objective underlies male involvement programs, and in some cases, they have no discernable behavioral objective. Rather, such programs view participation of males as an end in itself.

To the extent that male initiatives have had specific behavioral objectives, they have evolved over time, consistent with the four approaches outlined above. Under the first two (traditional and men and family planning), the primary objective tended to be the same as for women's programs: to increase contraceptive prevalence. However, programs that have adopted the third approach (male equity) seek to increase utilization of male-oriented RH services, the first of the objectives described below. Programs with a gender equity focus often have one or more of the remaining objectives listed below: to increase couple communication, to increase support for women's RH practice, and to change societal gender norms that harm women's health.

With the expansion of reproductive health beyond family planning to a broader range of subjects, the potential for male involvement also increases. For example, males often play the dominant role in sexual decision-making, directly related to prevention behaviors for HIV/AIDS. Men play a crucial role in providing financial and logistic support for women in need of emergency obstetrical care, as well as postabortion care. Adolescent programs seek to reach males as well as females, not only with information and services to protect themselves and their partners, but also with messages regarding gender equity. Males are also part of the intended audiences for initiatives to eliminate FGC. Whereas family planning programs tended to focus largely on women, the expanded range of RH interventions calls for an expanded role for men. In short, where the objectives of male initiatives extend beyond simply increasing contraceptive use, they tend to include one or more of the following objectives.

(1) To increase utilization of male-oriented RH services

Men's reproductive health needs include a wide range of services: family planning, treatment and prevention of STI/HIV/AIDS, infertility, sexual problems (impotency), and others. They also need clinics and doctors that provide confidential and non-judgmental care. Program staff should be aware that, like women, men are not a homogenous group. The needs of adolescent males, married men, older men, men with HIV/AIDS, and homosexual men differ. However, reaching men is more difficult than reaching women, for whom maternal and child health services have been designed. Traditionally, men are much less likely to have used a health clinic than women are.

Male-oriented services take several forms. Some countries have established male-only clinics. Others have established special hours (or evenings) for men at facilities oriented to women or to the general public. Others have publicized existing services more widely (e.g., STI treatment or vasectomy) to reach a larger number of men, or they have added services to increase options for males and to improve existing services (e.g., no-scalpel vasectomies).

(2) To increase couple communication

Research to date in the area of family planning shows that couples who make contraceptive decisions together tend to have higher rates of contraceptive use (Huezo and Malhotra, 1993). Women may have perceptions about their partner's ideal family size or desire for more children that prove false when the couple actually discusses the subject. Similarly, couple communication is essential for negotiating condom use or use of natural family planning.

While increasing couple communication is a desirable objective, we mention several caveats. Increasing couple communication does not necessarily improve communication. Nor does it ensure that women will have an equitable role in decision-making. Programs must ensure that they do not inadvertently undermine women's ability to make RH decisions by involving men. For example, one campaign in Zimbabwe used athletes to encourage men to play a greater role in family planning decision-making but actually increased the percentage of men who thought they should have sole control over contraceptive decision-making (Piotrow et al., 1992).

(3) To increase partner support for women's RH practice

In addition to addressing men's own reproductive health concerns, involving men may improve women's health. Women generally do not make decisions about their own reproductive health in a vacuum. Rather, their husbands as well as other family members influence them. Men are instrumental in providing the necessary physical, financial, and emotional support to give women access to reproductive healthcare, or conversely represent obstacles for many women who wish to protect their reproductive health. Some programs seeking to better women's health encourage men to:

- Ensure that their partners have adequate nutrition during pregnancy;
- Ensure that their partners seek prenatal care;
- Ensure that their partners seek emergency obstetrical care, if needed;
- Prevent the spread of STI/HIV; and
- Provide support for families by encouraging education for both boys and girls, adequate nutrition, and healthcare for all.

However, not all women want their partners "involved." Some use contraceptives against their husbands' wishes or without their husbands' knowledge. Others are afraid to raise the subject of family planning or condom use for fear of violence. In short, programs need to ensure that the female clients **want** their partners involved before programs proceed to involve them.

(4) To change societal gender norms that harm women's health.

Although changing social norms may seem an impossible task, over the past 30 years, we have witnessed a dramatic change in the norms governing the acceptability of family planning. The concept that having a large number of male children is a sign of masculinity has changed in many countries around the world. Some programs have begun to focus on men with the objective of influencing social norms related to promiscuity, forced sexual intercourse, and violence. In Nicaragua, for example, a UNFPA program "Apoyo a los Servicios y Acciones de IEC del Ejército de Nicaragua en Derechos y Salud Sexual Reproductiva" (Support for Sexual and Reproductive Health Services and IEC Initiatives in the Nicaraguan Army) worked to sensitize soldiers in the National Army to issues of sexuality, in particular STI prevention and gender-based violence (UNFPA, 2000).

Other programs have promoted specific behaviors that reduce risk to both men and women: encouraging abstinence or delay of sexual initiation, reducing the number of sexual partners, and consistently using condoms for sexual relations, especially with non-regular partners.

Programs dealing with larger cultural issues, such as female genital cutting and the early marriage of females, have also begun to include men in their programming with the objective of changing their attitudes on these deeply entrenched practices.

Methodological Challenges of Evaluating Male Involvement Initiatives

Despite the recent surge of interest in this area, relatively few organizations have evaluated effectiveness of these programs. Most projects have been small in scale, and few have been subjected to rigorous evaluation. Even though male involvement initiatives are still in their infancy, several challenges for evaluating these programs are already evident.

- **Many male involvement initiatives lack clear behavioral objectives.**

The purpose of program evaluation is to determine whether the program has achieved its objectives. In al-

most all other areas of RH, intermediate objectives refer to expected changes in knowledge, attitudes, and behaviors that, over time, lead to improved health outcomes (e.g., decreases in fertility, mortality, and morbidity). Many male involvement initiatives to date have focused on male participation in activities seemingly as an end in itself, not as a means to meaningful behavioral change. Others share the objectives of programs designed to reach women (e.g., increase contraceptive use), but lack the means to isolate the role of male involvement in the process. Programs generally work on certain underlying assumptions regarding the benefits of male participation, but few programs have designed interventions based on a clear conceptual framework – illustrating the pathways by which male participation will improve RH behaviors and outcomes – that provides the basis for systematic evaluation.

- **Because of the historic focus of the RH field on women, relatively little data exist on men.**

The 1980s saw growing recognition of the role men play in family planning decision-making; consequently, the DHS began to collect a limited amount of data on males, as a complement to the conventional DHS-data collection for women. With the spread of the AIDS epidemic, males have assumed a higher profile in RH interventions: as a result, the DHS and the RHS surveys have collected more data on men. However, routine service statistics often capture conventional service utilization, which focuses on women and children, not on men. In short, data on men are less available than data on women are.

- **Quantitative measures (from the DHS/RHS or service statistics) do not capture the complex issues involved in gender dynamics.**

Despite the traditional reliance on quantitative indicators, these indicators will not be sufficient to guide the development and refinement of male involvement programs. To complement the quantitative indicators (such as those listed in this *Compendium*), programs will want to use a variety of qualitative techniques to assist them in designing and evaluating programs in the future.

The indicators presented in this *Compendium* are by no means exhaustive. Rather, they represent illustrative indicators corresponding to the four common behavioral objectives of male involvement interventions, outlined above.

As mentioned above, the evaluation of most male involvement initiatives focuses on process measures of male participation in activities. Few to date have actually measured these interventions in terms of outcomes (e.g., knowledge, attitudes, behavior). The indicators that follow illustrate the types of measures that evaluators can apply to male involvement activities in the future, depending on the actual objectives of each intervention. In addition to the indicators presented in this section, male involvement programs can also use a number of the indicators presented for Behavior Change Communication Programs. (See Part II.F.)

NUMBER OF VISITS TO MALE-FOCUSED SERVICES, BY TYPE OF SERVICE

Definition

The number of "visits" – each occasion on which an individual seeks assistance from a given facility (or male-focused service within a larger facility)

The number of visits will be the same or greater than the number of persons using the service (reflecting some repeat visits to the service).

Data Requirement

Number of visits, by type of visit

Data Source(s)

Program service statistics

Purpose and Issues

The indicator reflects the volume of service provided to men for RH problems, as well as the nature of the problems treated. It is useful in justifying the continuation of this type of service if demand remains high. Also, it allows program managers to adjust staffing patterns based on the services most in demand.

An alternative indicator is the number of individuals seeking information and/or services. This indicator raises the question of possible repeat visitors, which is not an issue with "number of visits."

This indicator represents a minimum of information needed to track male use of RH facilities. The studies determining client satisfaction with the services will further enhance the monitoring process.

PERCENT OF MEN (HUSBANDS) WHO ARE SUPPORTIVE OF THEIR PARTNERS' REPRODUCTIVE HEALTH PRACTICES

Definition
The percent of males who support their partners' reproductive health practices

This indicator is calculated as:

> # of males who support their partners'
> reproductive health practices _____ x 100
> Total # of men surveyed

"Supportive" can be operationally defined in several different ways, including attitudes toward specific behaviors (e.g., contraceptive use), responses to hypothetical situations, and reported actions/behaviors.

"Reproductive health practices" refer to the behaviors that reproductive health programs promote (e.g., often the objective of the program): contraceptive use, breastfeeding, delivery in the presence of a skilled birth attendant, and so forth.

Data Requirements
Responses to structured or in-depth interviews

Data Source(s)
Surveys among the male clientele at health facilities or other men's reproductive health sites (program-based) or among the men in the general public (population-based). Alternative sources are surveys among the wives of participants in male-focused programs.

Purpose and Issues
One way a man can become "involved" in RH is by supporting his wife/partner in her practice of desirable health behaviors. Although some argue that this type of involvement does not go "far enough," in societies where males have withheld such support, this involvement can represent an important step forward.

Evaluators can assess men's level of support for women's RH practices using three types of questions: attitudes, responses to hypothetical situations, and reported actions. Illustrative questions of each type are presented in Box III.J.1. One expects that these responses will become more favorable as a result of interventions directed toward male involvement.

The answers to this set of questions are subject to bias, especially if men are aware that their attitudes or behaviors deviate from socially accepted responses. The best solution to this problem is for the interviewers to ask these questions in a matter-of-fact way. An alternative approach is to interview women about their husbands' attitudes and behaviors vis-à-vis family planning, safe pregnancy, delivery, STI/HIV risk, and other prevention behaviors. However, such accounts may be biased if the wives know that their husbands participate in the male-focused activities and thus "anticipate" changes in their behavior.

Box III.J.1 Illustrative Items for Measuring Men's Support of Their Wife's/Partner's RH Practices

Attitudes:

Do you approve or disapprove of your wife's/partner's:

(a)   Using a contraceptive method to prevent pregnancy?
(b)   Receiving antenatal care during pregnancy?
(c)   Having a trained birth attendant present at delivery?
(d)   Breastfeeding your baby?

Hypothetical situations:

1) If your wife/partner went into labor and experienced complications but you were away on a trip, should she seek health care on her own or wait for your return?
2) Suppose a woman suspected that her husband/partner was having sexual relations with several other women. Is she justified or not to suggest using condoms when she and her husband/partner have sex?

Actual behaviors:

1) Have you ever told (or otherwise let your wife/partner know) that you approve or disapprove of her using contraception?
2) During your wife's/partner's last pregnancy, did you have a plan to get her to a hospital or health center if she had complications? (If so, explain).

PERCENT OF CONTRACEPTIVE METHOD USE REQUIRING MALE COOPERATION

Definition

"Contraceptive method use" refers to the number of individuals in a population, clients in a program, or respondents on a survey who report using some type of contraceptive method (female or male). Methods requiring male cooperation include vasectomy, condoms, withdrawal, and periodic abstinence (e.g., rhythm).

This indicator is calculated as:

$$\frac{\text{\# of users reporting use of condom, vasectomy, periodic abstinence, or withdrawal}}{\text{Total \# of contraceptive users}} \times 100$$

Alternatively, this indicator may also be calculated for new adopters of contraceptives:

$$\frac{\text{\# of new adopters who opt for condoms, vasectomy, or periodic abstinence}}{\text{Total \# of new adopters}} \times 100$$

Data Requirements

Population-based responses (among married women of reproductive age); program-based on new adopters during a reference period (e.g., one year)

Data Source(s)

DHS, RHS, or other representative survey; service statistics

Purpose and Issues

This indicator measures the extent to which men take responsibility for contraception within their own marriage or other sexual union. Family planning programs have traditionally focused on women, and the large majority of contraceptive use consists of female-controlled methods (oral pill, IUD, injectables, implants, female sterilization).

One advantage of this indicator is that the data are readily available through routinely collected service statistics or through DHS or RHS surveys. The shortcomings are two-fold. First, male involvement interventions usually target a limited geographical area, in which case these large-scale surveys lack appropriate data for evaluation (although a representative survey of the area will). Second, program-based service statistics are readily available. However, they fail to capture contraceptive use outside the government or NGO facilities that provide family planning (e.g., pharmacies, which are a major source of supply of condoms).

Not all male involvement initiatives are designed to increase male responsibility for contraceptive use. Thus, this indicator is appropriate only when the intervention has this specific objective.

PERCENT OF MEN AND WOMEN WHO DISCUSS REPRODUCTIVE HEALTH ISSUES WITH THEIR SPOUSE OR SEXUAL PARTNER

Definition
The extent to which couples discuss RH issues

"Reproductive health issues" are operationally defined in relation to the local context (see illustrative questions below).

Data Requirements
Responses to interviews

Data Source(s)
DHS-type survey; special survey; and survey among clients

Purpose and Issues
Many societies prohibit men and women (even husbands and wives) from discussing RH issues, such as contraceptive use, condom use for STI prevention, women's nutritional needs during pregnancy, and so forth. Male involvement interventions often are designed to increase male awareness of RH issues and to increase partner communication on these topics. This indicator measures the extent to which husbands and wives or other sexual partners discuss specific RH topics. The concept of inter-partner communication is somewhat open-ended. Questions on partner communication need to be clear and concrete to foster valid responses. Illustrative questions include the following:

"In the past month, did you and your spouse/partner discuss:

- Using or continuing to use contraception to prevent pregnancy?
- Using condoms to prevent sexually transmitted infections, including HIV?

- The need for a skilled birth attendant at delivery (if pregnant)?
- The need for your partner to get adequate nutrition and rest (if pregnant)?
- Whether or not to circumcise your daughter (if appropriate)?"

A one-time measure of communication between partners is useful for diagnostic purposes, but not as an indicator for monitoring or evaluating a program. Rather, the evaluator must use some type of study design that shows the effects of the intervention on increased inter-partner communication.

Evaluators can design studies to:

- Compare the percentage of clients who report inter-partner communication on specific topics before and after counseling or other BCC intervention (note: this approach requires interviewing the client on two separate occasions); and
- Compare the percentage of clients in experimental versus control groups who report inter-partner communication on specific topics (i.e., in the former group, the male partners participated in the counseling session, in the latter group they did not).[2]

---

[2] This design requires random assignment of subjects to the experimental vs. control group, as well as a before/after measurement.

Part III.K
Violence against Women

- Existence of a policy on violence against women
- Number of VAW service visits provided, by type of service
- Attitudes of health care providers towards VAW or VAW services
- Percent of clients satisfied with the VAW service on multiple dimensions

Violence against women (VAW)[1] is a serious health and social problem that jeopardizes women's reproductive health and violates women's reproductive rights. Cross-national research into the scope and nature of violence against women has found very high prevalence rates of VAW and a strong relationship between VAW and adverse reproductive health (Heise, Ellsberg, and Gottemoeller, 1999). Women who experience forced or unwanted sex are at increased risk of STIs including HIV, and are less able to engage in behaviors that will protect them from STIs and HIV. Sexual dysfunction, poor mental health, adverse pregnancy outcomes, and drug and alcohol use are all more common among women who have experienced sexual and/or physical violence. The health sector has a critical role to play in the response to VAW, because it provides the only service that all women are likely to receive at some point in their life. Because violence increases the risk of many other health problems, identifying it early may help prevent other serious and life-threatening diseases (Heise, Ellsberg, and Gottemoeller, 1999).

Social science research indicates that VAW is a socially normative behavior in many societies (Sanday, 1981; Heise, Ellsberg, and Gottemoeller, 1999) and that it functions to achieve both short-term goals of social interaction and longer-term goals of maintaining female subservience and male dominance (Dobash et al., 1992). Thus, many societies may consider it "normal" for a husband to use physical force against his wife to "keep her in line"; other common excuses for VAW, domestic violence in particular, include real or imagined extra-relationship sexual activity or transgressions of expected sex-role behavior, such as performing household tasks. This violence can range from slapping, pushing, or shaking to strangling, burning, threatening with a weapon, or rape. Violence against women may take specific forms in specific countries, and thus we must understand how the local context conditions both the forms and consequences of VAW.

This section of the *Compendium* focuses on violence against women in the context of reproductive health program evaluation. In contrast to several other areas of reproductive health in which mature programs are underway at the national level to combat a given problem or to promote specific behaviors, programs on violence against women are still in their infancy. VAW services tend to be organized in one of two ways in developing countries. In a limited number of countries, stand-alone organizations (generally NGOs) primarily function to combat violence against women. Their programs often include a diverse array of activities and services (e.g., advocacy to legislators, communication programs through the mass media and interpersonal channels, onsite legal advocacy services, health and social services for victims of VAW, and related activities). In such cases, reproductive health is one part of a larger program of activities. Other countries do not have stand-alone programs but rather offer some assistance to victims of VAW as part of a larger range of health or social services. Such organizations tend to provide counseling and treatment for medical conditions and, in some cases, psychological trauma, but make referrals to other non-health services (e.g., legal). All of the indicators in this section apply to the stand-alone facility (assuming it gives health care or social services). By contrast, the indicator on the existence of a policy is much more relevant to a stand-alone program with an advocacy agenda than to an RH program offering VAW services among numerous other services.

Depending on the mission of the organization, the VAW programs have one or more of the following objectives:

---

[1]Gender-based violence is an umbrella term for any harm that is perpetrated on a person against her/his will, and jeopardizes the physical and/or psychological health, development, and identity of the person. Violence results from gendered power relationships, determined by the social roles ascribed to males and females, disproportionately impacting women and children in almost all cultures. Violence may be physical, sexual, psychological, economic, or socio-cultural. Categories of perpetrators may include family members, community members, and/or those acting on behalf of cultural, religious, or state institutions.

- Developing national health and social policies on the issue;
- Sensitizing health care providers to VAW and training them in techniques to recognize appropriately and to respond compassionately to VAW;
- Training staff in specific protocols and procedures for VAW: collecting forensic data, conducting danger assessments, creating safety plans;
- Providing information/raising awareness within a clinic or hospital setting via educational materials;
- Identifying women who have experienced VAW via patient screening;
- Referring women to internal or external services (legal advocacy, shelter services, psychological support, among others); and
- Providing direct services such as legal advocacy or psychological counseling.

We have developed several indicators that reflect the activities related to policy development and service provision. In terms of the latter, we have included both a quantitative indicator on volume of services provided and a more qualitative assessment of client perceptions of the services. We have omitted population-based VAW indicators (even though decreasing violence against women is the ultimate objective) for the following reasons.

First, in the case of national-level programs such as family planning, one may expect to see the effects of the program at the population level (e.g., an increase in contraceptive use among married women of reproductive age). However, in programs aimed at a small subset of the population or having a limited reach (e.g., adolescent youth centers), we cannot expect to detect changes attributable to the program in population-level data. In programs like these, use of program-level data (e.g., information obtained from clients who come into contact with the program) is more appropriate.

Second, obtaining valid data on VAW from population-based surveys requires careful attention to a series of methodological and ethical issues. The WHO has developed a set of guidelines for conducting research on VAW with the goal of enhancing the safety of respondents and interviewers and of obtaining valid data. The most critical point is that many countries lack the capacity to address the VAW issues appropriately (once identified through research), given the dearth of official mechanisms or social services designed around VAW.

Third, research conducted in the absence of such measures often results in much smaller prevalence estimates than expected, based on similar population-based studies, thus causing policymakers to think that there is "no problem." For these three reasons, we do not advocate using prevalence of VAW as a standard indicator of program effectiveness and have included instead an indicator on the existence of national- or institutional-level policies. Given the desirability of having some population-based measures, one alternative is to track attitudes at the community level concerning the justifiability of VAW, the forms it takes, and the perceived frequency of occurrence.

Finally, the goal of reducing violence against women is laudable but long-term. Achieving this goal through interventions by the health care system *alone* is unlikely. If evaluators use prevalence of VAW to measure "success" of an individual program, they may erroneously conclude that the program has "failed."

Macro International Inc. has developed a module on domestic violence for use in connection with the Core DHS questionnaire. This module includes a series of questions on abusive or violent behavior on the part of the (last) husband/partner, when it started, what physical consequences occurred, whether he drinks excessively, and so forth. The module also includes questions on possible violent behavior of the woman toward her partner, on violent behaviors from other family members and acquaintances, violent behavior of the partner during pregnancy, help sought, reasons for not seeking help, and related topics. The primary purpose of the module is to determine the extent of domestic violence in a given society, the most common forms of violence, the context of these violent actions, the perpetrators, and related information. To date, no country has tried to use such data to evaluate program initiatives for eliminating violence.

The RHS has a module on Intimate Partner Violence (IPV) used in all surveys conducted in Eastern Europe since 1997 and some surveys in Latin America since 1995. The questions focus principally on two types of violence against women: 1) intimate partner violence during the respondent's lifetime and within the past year;

and 2) sexual coercion at any point in a woman's life. Violence by an intimate partner – defined as verbal, physical and sexual abuse – is explored using a modified Conflict Tactic Scale (eight items). In addition, all respondents are asked about their history of witnessing physical abuse between parents or their experience of abuse as a child or adolescent. These data are powerful for advocacy purposes. An NGO and Senator in Paraguay used data from the 1995 RHS in Paraguay to get the first law against violence against women passed in that country.

Methodological Challenges of Evaluating VAW Programs

Although very little empirical work has been done to evaluate the effectiveness of programs designed to reduce violence against women in the international context, we have included this section on VAW in the *Compendium* to stimulate discussion of the most appropriate measures to use, in anticipation that more programs of this type will be implemented in the future. Despite the dearth of empirical data on these programs, researchers have already identified the following as methodological limitations of conducting work on this topic.

- **Program-based statistics on volume of clients are difficult to interpret.**

Often programs may identify reducing domestic violence as a long-term objective, while offering services to women in abusive situations in the short-term. Statistics generated by such programs are often used to describe the nature and scope of the problem of domestic violence; however, service utilization statistics are open to various interpretations and may not be an appropriate measure of the occurrence of violence within a certain area or among a certain population. If for example, the number of women enrolling in a domestic violence program increases, one might conclude that the incidence of domestic violence is actually on the rise. In fact, the change may be due to a greater awareness among women of the availability of such services and a willingness to try them, based on information they have received (e.g., favorable word-of-mouth, mass media messages) or an increase in resistance by women to men's use of violence against them. As such, the increase in the number of women using the services has an ambiguous meaning.[2] This difficulty of interpreting program data tends to be greatest during the early years of the program.

- **Quantitative program statistics do not reflect the quality of the service provided.**

Although a program may document that it has provided 400 units of service to female clients, quantitative data fail to reflect the quality of services provided and its effectiveness in empowering the client to face her situation. For example, a program may distribute a large volume of materials, but this activity may do more harm than good if the material carries a victim-blaming message. If health providers are callous and insensitive in screening women for intimate partner violence, then the benefit of this service to beneficiaries is questionable at best. Similarly, if providers screen for violence but do nothing for clients who are victims, the statistics on "numbers screened" are meaningless. Finally, if clients perceive the referral process simply to be a means for providers to get rid of them, then "number of referrals" inadequately measures program performance. Qualitative data that reflect clients' perceptions of the interaction with the service providers is essential to complement the quantitative objective indicators proposed in this section.

- **Identifying clients who received VAW support services for interview may endanger them and the interviewer.**

VAW program evaluation faces a further difficulty: identifying and interviewing women who received such services may place the woman at further risk. Evaluators must plan and execute data collection with utmost care to protect the client. In addition, the interview may place the interviewer in a position where she is ethically or morally obliged to intervene (e.g., if the respondent indicates that she may be in immediate lethal danger). In program evaluation, one discourages the interviewer from "switching roles" at the end of the interview and becoming a health educator or counselor. (Among other reasons, the interviewers are not trained to provide information or counseling, and their well-intentioned responses may be incorrect.) Collecting this information and doing nothing for the woman in this potentially dangerous situation would be unethical. Given the sensitive nature of the issue and the potential for endangering the client, only programs or program evaluators skilled in VAW research techniques should collect data directly

---

[2] This problem is not unique to VAW. For example, a rise in treatment-seeking for STIs following a campaign on the topic does not necessarily (and even likely) indicate a rise in STI incidence.

from program clients. Readers are referred to the WHO Guidelines (1999d) for further information on the topic.

- **Obtaining valid data on provider attitudes is difficult.**

Because of the difficulties involved in obtaining data from women who have experienced violence, a tempting alternative is to interview providers instead. However, attitudinal data may be of limited value, because (in VAW as in other areas) attitudes may not be predictive or strongly correlated with behavior. Maiuro et al. (2000) have published a useful article on the subject of measuring provider attitudes towards VAW within the U.S. health care system, though the results may not generalize to developing countries.

The following pages contain several indicators for evaluating initiatives to combat VAW. In contrast to other sections of the *Compendium* that focus on population-based outcome measures, this section focuses instead on policy issues and program-based measures. As outlined above, it is premature to think of population-based measures for the evaluation of programs to combat violence against women, given the very small-scale on which such programs are implemented, not to mention the ethical issues involved with maintaining "scientific rigor" to the possible detriment of the subjects involved in such evaluation work.

Readers interested in pursuing this topic in greater detail are referred to a research brief entitled "Reproductive Health Program Responses to Gender-Based Violence Against Women: Conceptualizing Indicators for Monitoring and Evaluation" (Frye, Banwell, and Ellsberg, 2001). This paper further develops some of the ideas outlined herein (e.g., in regard to knowledge and practice indicators for providers, administrators, and managers).

EXISTENCE OF A POLICY ON VIOLENCE AGAINST WOMEN

Definition
The existence of formal governmental declarations, laws, and statutes affecting VAW Policy also can refer to operational regulations, guidelines, norms, and standards (Cross, Jewell, and Hardee, 2001).

Data Requirements
Documentation outlining the policy

Data Source(s)
Legislative records, administrative records, and other government documents (national, regional, and local); also, internal policy documents of an organization

Purpose and Issues
The purpose of this indicator is to track changes in the policy environment that potentially affect the delivery of VAW services and the well-being of victims of VAW. Such changes can occur in the political arena (via formal governmental declarations and changes in legislature, which some refer to as Policy with a capital "P") or at the organizational level (in terms of the policies and procedures used within RH health services and by institutions that refer women to these services, such as the police, judiciary, and social services (policy with a small "p"). This indicator is a specific case of the indicator on **Existence of Policies, Plans, Guidelines that Promote Access to and/or Quality of RH Services** presented in Part II.B.

Experts in this field maintain that any organization dealing with VAW should article a policy (small "p") on its approach to gender-based violence. The organization may also take an advocacy stance and try to influence governmental policy and legislation (capital "P"), depending on its mission.

In analyzing policy-related documents on VAW, one should further consider how the document frames the issue:

- Does the policy acknowledge VAW as a complex and multi-determined social and health phenomenon?
- Does the policy use an integrated approach to respond to VAW or a fragmented, single-sector approach?
- Does the policy work collaboratively with women's organizations that have been on the front line of the response to VAW up until the point of a formal policy development?

Whereas the existence of a policy on VAW signals political concern over the topic, it may be relatively meaningless if not translated into concrete actions. Any assessment of VAW policy should examine the actual structures in place to respond to the needs of victims of VAW, as well as the record of implementing the policy initiatives to prevent violence in the future. Thus, a related indicator involves the existence of structures to (1) provide services to those who experience VAW and (2) undertake initiatives aimed at reducing or eliminating VAW in the future.

NUMBER OF VAW SERVICE VISITS PROVIDED, BY TYPE OF SERVICE

Definition

"Service visits" are counted as the number of occasions on which a woman seeks VAW assistance from a given center. The total number of visits may include repeat visits and thus may be larger than the total number of women using the center or program in a given year.

Note: A woman may receive more than one service on a given visit (counseling plus referral for other health problems). Program managers and evaluators may find it useful to track the different types of services (e.g., counseling, screening, referrals, treatment for injuries from violence) to better understand the needs of the clientele. For example, this tracking would yield data on the number of referrals made from VAW centers to related services in the course of a reference period (e.g., one year).

Data Requirements

Number of visits per center, aggregated across multiple centers (if such exist)

Data Source(s)

Service statistics from the center or program

Purpose and Issues

This indicator measures the volume of services the program provides to its clientele. During the early years of the program, evaluators should monitor details regarding the visits to better understand the problems and potential needs of the clientele (e.g., reason for the visit, type[s] of services provided).

Several related indicators (for RH facilities) include the following:

- Number of clients reporting violence as a percentage of all women seeking RH services;
- Active screening for VAW: percent of the total VAW reports that were identified through active screening; and

- Timely and appropriate post-rape care: percent of survivors who access service within three days of assault.

Clients coming to the facility for other services are more likely to divulge an episode of violence if they sense providers will be sensitive to their problem.

As noted in the introduction to this section, the interpretation of this indicator is somewhat ambiguous. The number of visits could increase over time, not because violence against women is mounting, but rather because women are more willing to come forward and disclose this problem, especially if the word-of-mouth information about the center is favorable. In fact, an increase in service delivery should reflect favorably upon the program.

This information is useful to demonstrate to donor agencies that the organization is providing a service within the community. Again, the indicator gives little sense of whether the women who receive the service perceive it to be helpful, although an increase in numbers may reflect favorable word-of-mouth publicity. Also, the number should rise as a result of mass media publicity or other BCC interventions on VAW. Ideally, the statistics on number of visits will also rise, especially during the early years of the program, as more women in need learn that services are available and helpful to women who experience violence.

Although the program may not be able to demonstrate effects at the population level, data on service utilization will help justify the continued existence of the services to donors interested in assisting women with the problem of domestic violence.

ATTITUDES OF HEALTH CARE PROVIDERS TOWARDS VAW OR VAW SERVICES

Definition

The attitudes of service providers towards women's socially prescribed sex-roles, the issue of VAW, the VAW service they provide, and the women who receive the services

Attitude is defined as a person's favorable or unfavorable assessment of a behavior or situation.

Data Requirements

Responses to surveys; transcripts from focus groups

Data Source(s)

Interviews of service providers; and focus groups

Purpose and Issues

The indicator identifies providers who hold victim-blaming, fatalistic, passive or other attitudes inconsistent with gender-sensitive quality of care. Illustrative examples of attitudes to measure are presented in Box III.K.1. These individual-level provider attitudes are important to track because they constitute barriers to (1) women's reporting and seeking VAW services and (2) the delivery of sensitive and appropriate services.

This information demonstrates to donor agencies that the organization is providing the service compassionately and sensitively within the community. In addition, the indicator reflects the quality of training that the organization provides to its health care providers.

Box III.K.1 Illustrative Attitudes to be Assessed among Health Care Workers for VAW

Sex-stereotyping:

- A woman must be a virgin when she marries;
- A wife should never contradict her husband;
- It is acceptable for women to have a career, but marriage and family should come first; and
- There is something wrong with a woman who does not want to marry and raise children.

Acceptance of interpersonal violence:

- Being roughed-up is sexually stimulating and/or a sign of a man's love for a woman;
- Women will pretend that they do not want to have intercourse because they do not want to seem loose, but they are really hoping the man will force them;
- A wife should move out of the house if her husband hits her; and
- A man is sometimes justified in hitting his wife.

Source: Burt (1980).

PERCENT OF CLIENTS SATISFIED WITH THE VAW SERVICE ON MULTIPLE DIMENSIONS

Definition
The degree to which clients report satisfaction with various aspects of the services received.

This indicator is calculated as:

$$\frac{\text{\# of clients who report satisfaction with the services received}}{\text{Total \# of respondents}} \times 100$$

Data Requirements
Attitudinal responses (interviews) or verbatim text (focus groups) on different aspects of services received (see Box III.K.2)

Data Source(s)
In-depth interviews or surveys with clients; focus groups conducted with clients

Purpose and Issues
The clients' subjective perception of the service received is a critical element of the program evaluation. It allows the evaluator to assess the extent to which the program services meet the needs of its clientele. If the clients do not perceive the service as helpful or if they feel that it further frustrates or disempowers them, the intervention has failed. In addition, client feedback is critical to the process of identifying ways to improve the program. Finally, evaluators can compare feedback from clients to information from providers to arrive at a more balanced understanding of program dynamics.

This indicator is prone to courtesy bias on the part of respondents. The pervasive tendency of respondents in exit interviews to respond positively about the services they receive (more positively than they feel about the services) is widespread among clients of any services.

However, this tendency is probably more prevalent among women who experience domestic violence because they feel especially vulnerable and fear that any negative feedback may affect their future access to these services. Evaluators should particularly note any item that clients rate even slightly lower than others as indicative of possible areas in need of improvement. Also, because the likelihood of frank responses to this series of questions may differ from one country to another, this information is most useful in evaluating a given center over time, rather than in making cross-national comparisons.

Here the ethical issues around research with women who experience VAW is important (note: see discussion in introductory section). Representatives of local women's organizations may be useful in ensuring that data collection does not further jeopardize the client question.

This indicator on client satisfaction measures one aspect of quality of care. An alternative approach, which avoids the problem of courtesy bias, is to measure certain objective aspects of the service delivery environment: staff training, availability of private screening rooms, and existence of protocols and their implementation. Although the Service Provision Assessment (the facility-based survey available for use with the household DHS) does not address VAW *per se*, evaluators may adapt many of the items for this purpose.

Box III.K.2 Illustrative Attitudinal Questions on Client Satisfaction with VAW Services

Client Interview:

- Did the health care provider (HCP) assure you of privacy and confidentiality?
- Did the HCP respond to your concerns and meet your needs?
- In your opinion, did the HCP address your concerns and needs?
- Did you feel as if the HCP treated you with respect?
- Did you feel that the HCP listened to you?
- Were you able to articulate your concerns and needs concerning violence to the HCP?
- Did you feel as if the HCP was trying to force you into a certain decision regarding VAW?
- Did the HCP provide information about your health, social, and legal options in a way that you could understand?
- Did you understand your health, social, and legal options?
- Did the HCP give you concrete information that helped you make your decision?
- Did you feel that the HCP wanted you to make your own decisions?
- Did you feel that the HCP supported the decision you made?
- Did the HCP make you feel as if you were to blame for the VAW situation?
- *If received referrals*, Did the HCP seem to believe that the referrals provided would help you?
- *If received referrals*, Did the HCP provide these referrals in a way that showed they cared?

(Banwell, Ellsberg, and Frye, 2001)

- Percent of population that know about FGC: legal status, religious position, health risks
- Percent of the population favorable to the continuation of FGC
- Percent of women 15-19 years old who have undergone female circumcision

Female genital cutting (FGC) or female circumcision is a generic term for traditional practices involving the cutting of female genitalia leading to the partial removal of the female genitalia or injury to the female genital organ for cultural or any other non-therapeutic reasons (Toubia, 1995; WHO, 1995d). Another term for this practice is female genital mutilation, which emphasizes the permanent physical damage done to the female genitalia (Yoder, Camara, and Soumaoro, 1999).

Female circumcision is deeply rooted in many African societies. The practice occurs in over 16 countries, but is rare or unknown in at least 20 other African nations. National borders are less relevant to delineating zones for this practice than are transnational cultural zones (Akweongo et al., 2001). More a secular than a religious phenomenon, FGC is found in both Muslim and Christian societies. Although a number of countries have banned female circumcision, the degree of enforcement varies from one country to another. For example, in Guinea circumcision is punishable by life imprisonment, but no one has ever been indicted for this crime (Yoder, Camara, and Soumaoro, 1999). By contrast, in Ghana authorities make every effort to prosecute practitioners who are caught; the media widely publicize these cases and identify perpetrators, as a form of social humiliation intended to deter others (Reason, 2001).

The WHO (1996d) has classified four types of female circumcision:

- Type I: Excision of the prepuce with or without excision of part or all of the clitoris;
- Type II: Excision of the prepuce and clitoris together with the partial or total excision of the labia minora. This type accounts for 80 percent of all cases;
- Type III (infibulation): Excision of part or all of the external genitalia and stitching/narrowing of the vaginal opening. Infibulation is mostly found in North Africa; and

- Type IV: Includes enlargement of the vagina (introcision); the pricking, piercing, incising, or cauterizing of the clitoris; the scraping of surrounding tissue of the vaginal orifice; or the cutting of the anterior, posterior vaginal wall (gishiri cuts); and sometimes the introduction of corrosive substances or herbs into the vagina to cause bleeding or for the purpose of tightening or narrowing.

According to Obermeyer's review of over 400 articles and reports on FGC published since 1995, the consequences of FGC include: (1) short-term effects, such as pain, hemorrhage, shock, and infection; (2) long-term effects, such as urinary infection, scarred tissue, fertility problems, and complications during child birth; and (3) long-term effects on the woman's sexuality and her social and affective relationships (Obermeyer, 1999). Although many assume that female genital cutting often results in death or severe complications, Obermeyer's review provides little evidence to support this contention. Rather, no incontrovertible evidence on mortality exists, and the available research suggests that severe complications are relatively infrequent. Political, economic, and ethical factors at both the local and international levels may explain this apparent contradiction between popular assumptions and available empirical evidence. Obermeyer concludes, "the scarcity of evidence regarding the complication of female genital surgeries is probably due to the lack of concerted efforts to investigate harmful effects rather than to the relative safety of these operations."

The response of the international community to FGC has swung from the cultural absolutism of Christian missionaries in the first half of the twentieth century, to a cultural relativism, and more recently back to the absolutist stand advocated by most sections of the Women's Movement and increasingly by international institutions (Caldwell, Orubuloye, and Caldwell, 1999). In the 1920s and 1930s, Christian missionaries in Kenya attacked the

practice on the grounds that it conflicted with Christian mores. For the next 50 years, the prevalent view was one of cultural relativism. In the 1950s, WHO avoided taking a stand against FGC on the grounds that these were "operations based on social and cultural backgrounds." By contrast, over the past 20 years, the women's movement has collaborated with NGOs and brought pressure on international organizations to work for the elimination of FGC. The practice remains bewildering and abhorrent to many Westerners, who fail to comprehend the reasons this practice should exist.

The motivations behind FGC are complex. Feminist groups have attributed the perpetuation of the practice to African traditions of male dominance and of the patriarchal system. Those who support FGC believe that it purifies the girl (by reducing her sexual desire), favorably socializes her through the instruction and training she receives during her seclusion, and ensures fidelity. One widely held view in some countries is that men prefer to marry circumcised women and will pay more in brides' wealth for them, although this is by no means consistent over countries. Caldwell, Orubuloye, and Caldwell (2000) cite respect for tradition and social conformity: "the central issues are fears of making their daughters seem outside the expectations of society and possibly unmarriageable, and making themselves also the objects of deep suspicion."

Two aspects of FGC absent from portrayals of this practice in the Western media are (1) that women play a key role in sustaining the practice, and (2) that, in some societies, the girls "decide" whether to undergo FGC (Akweongo et al., 2001; Yoder, Camara, and Soumaoro, 1999; Caldwell, Orubuloye, and Caldwell, 2000). Traditionally, older women (including mothers, co-wives, and heads of compounds) sustain the practice by exerting enormous pressure on young girls to undergo the procedure. Social ostracism and mockery rather than physical coercion are often used to ensure that the girl gets circumcised.

Several studies to date indicate that although the practice remains deeply rooted, the seeds of change are evident among more educated, urbanized populations. In a focus group study in Northern Ghana, the predominant view still favored FGC. However, a minority believed that the negative messages once directed to the uncircumcised are now more typically expressed as negative attitudes toward the practice (Akweongo et al., 2001). In one area of Guinea, women did not seem to want to abandon the practice, but they are ready to adopt a less severe form of FGC (Yoder, Camara, and Soumaoro, 1999).

In most countries where FGC is practiced, local groups (often NGOs) bolstered by international supporters have developed programs to combat FGC. Four intervention strategies used to reduce the practice of FGC[1] include: (1) raising awareness, (2) selecting some members of the community to serve as change agents (facilitators) in their communities including individuals who have resisted FGC (positive deviants), (3) integrating anti-FGC messages into development activities, and (4) strengthening advocacy (Abdel-Tawab and Hegazi, 2000).

The evaluation of such initiatives should ideally link to a model of behavior change. Specifically, evaluators can then select indicators to determine progress toward the desired outcome. Izett and Toubia (1999) describe five stages of behavior change in relation to FGC: precontemplation, contemplation, preparation, action, and maintenance. This model overlaps somewhat with the ideational change model presented in Part II.F. The indicators presented in this section reflect the components of knowledge, attitude, and behavior of the latter model. However, one important difference is that these three outcome indicators may each refer to a different player in the decision for FGC. For example, we want to measure knowledge and attitudes related to FGC among those who will decide or encourage a young girl/woman to have the procedure done (e.g., village elders, mother, mother-in-law). However, measurement of the practice – the outcome in this case – is based on the young woman who has (or has not) undergone FGC.

To date, evaluation of anti-FGC interventions has been minimal. To the extent they exist, such efforts have focused on process: participant turnout to seminars and other events, participant reaction to the seminar, and knowledge gain measured by pre- and post-tests. The observation of Abdel-Tawab and Hegazi (2000) relative to Egypt equally applies to most countries where FGC is practiced:

> NGOs seldom document the process they follow in implementing interventions, the strengths and weaknesses of each approach, difficulties faced in implementation, or ways of overcoming those difficulties.

---

[1] Although the authors based their analysis on Egypt only, these same categories appear to apply to other countries as well.

Also, there is scant empirical evidence about the impact of these various models of programmatic interventions.

One promising source of good data on FGC is the module developed by Macro International in connection with the DHS survey. To date, this research has been diagnostic in nature, conducted to better understand the extent of the practice and the conditions under which it occurs. The three indicators presented in this chapter are all population-based, and two of the three are available through the DHS module. Evaluators can use them to track change, though in most cases, they will have difficulty attributing change uniquely to program interventions (in the absence of a control group). These same population-based indicators can serve in smaller scale studies designed for the express purpose of evaluating FGC interventions. Since eradication of this practice may take years to achieve (despite notable progress in some countries such as northern Ghana), evaluators should track evidence of change in the form of knowledge and attitudes, in addition to the actual reduction in practice of FGC.

Methodological Challenges of Evaluating Programs to Eradicate FGC

To date, governments and NGOs have tried different approaches for eradicating FGC. In the 1980s and 1990s, advocacy groups exposed the practice in selected countries through the mass media, in the hopes that the international community would exert pressure on local governments to ban the practice. Indeed, FGC is now illegal in numerous African countries: Burkina Faso, Central African Republic, Ivory Coast, Djibouti, Ghana, Guinea, Senegal, Tanzania and Togo (The Center for Reproductive Law and Policy - CRLP, 2001). However, these countries may or may not enforce the laws. A second wave of initiatives, beginning in the 1990s, has attempted to eliminate FGC by helping communities to understand the factors that sustain FGC and to explore alternative strategies for ushering girls into womanhood. These initiatives seek to conserve the positive cultural values associated with the traditional ceremonies, while eliminating the physical and psychological trauma of FGC (Nazzar et al., 2001; LSC, 1998a; LSC, 1998b). Few organizations have systematically evaluated these initiatives; often implementation is sufficiently daunting that organizations do not even consider evaluation. However, notable exceptions are now emerging.

For example, the Navrongo Health Research Centre is conducting a community-informed experiment in preventing female genital cutting among the Kassena-Nankana of northern Ghana. During the first phase of this four-year experiment, the research team will use qualitative data research methods to clarify the complex social rationale behind the practice of female circumcision, to identify socially acceptable strategies of responding to these traditions, and to identify outreach activities for preventing FGC. In the second phase, the team will use lessons learned from the village micropilot to test the impact of the Phase I strategy on a larger scale in the community. The intervention will continue for a period of four years. In contrast to many of the national-level programmatic interventions covered in this *Compendium*, this project will experiment in a single, small, isolated village in northern Ghana. This project illustrates two of the potential methodological problems that can emerge in evaluating interventions to eliminate FGC:

- **As people become increasingly aware that these practices are illegal and socially unacceptable, response bias will increase.**

As programs to prevent these practices reach an increasing number of people, those who may previously have reported the practice will become increasingly reluctant to do so. One approach to combating this problem is to obtain information from more than one source (e.g., the young woman, her parents, and other community members).

The incidence of underreporting may relate to age of the respondent, especially if younger women are more aware of the anti-FGC initiatives and/or are more motivated to appear "modern." Thus, comparison of percentage circumcised by different age cohorts may be subject to this bias.

- **Members of the key population may leave home, creating a problem of "censoring" in the data.**

In the case of FGC research, a key population of interest is young women. However, young adults often leave their rural settings to pursue economic activities in larger cities. In areas with high levels of migration toward urban areas, studies in rural areas may have a considerable "lost to observation" rate for adolescent women

(Nazzar et al., 2001). Results will be biased if those who migrate are less likely to be circumcised than those who stay are (i.e., selectivity).

A third methodological consideration, addressed in connection with the 1995 Egypt DHS, is the following:

- **Women may not be able to accurately report if they are circumcised or not.**

Self-reported data are always subject to bias, especially in relation to a medical procedure such as the type of circumcision performed. Some FGC researchers have questioned whether women **know** whether they are circumcised; even their husbands may not know for sure.

This question arose in connection with the 1995 DHS in Egypt, a country with high prevalence of FGC (97 percent as of 1995). A special clinic-based study compared the clients' responses (self-report) to physical evidence obtained at the time of a pelvic exam performed by specially trained gynecologists. The 1,339 women included in the study – clients at the clinic for family planning or gynecological problems – were not representative of the national population, but provided a useful basis for this assessment. In 94 percent of the cases, the woman's self report coincided with the physical evidence of the amount of tissue excised during circumcision. In 5 percent of the cases, the women reported circumcision when in fact the gynecologists found no physical evidence of it. And one percent of women reported that they were not circumcised, when in fact they were (El-Zanaty et al., 1996).

These findings from this one study suggest that women are able to reliably report the type of procedure performed. However, these findings conflict with anecdotal evidence that some women may not even know if they are circumcised, let alone the type of circumcision performed. Moreover, as promotional/informational programs on FGC become more frequent and FGC becomes less socially acceptable or "modern," then the reliability of self-report may diminish.

The indicators that follow are all quantitative in nature. To develop a more thorough understanding of FGC, qualitative research is essential. One promising area for additional research relates to coming-of-age rituals and puberty rights. For example, as Reason (2001) has asked, "Are there changes over time in response to initiatives to eradicate FGC?" Because FGC is a new area for program evaluation, we limit the number of proposed indicators to three, with the expectation that as work in this area evolves, additional indicators may emerge. Earlier drafts of this volume contained a fourth indicator on the type of (severity of) circumcision performed, with the rationale that moving toward less radical forms of the procedure represented some type of progress. However, on the advice of reviewers, we dropped this indicator on the grounds that it may imply acceptance of less severe forms of FGC.

PERCENT OF THE POPULATION THAT KNOW ABOUT FGC: LEGAL STATUS, RELIGIOUS POSITION, HEALTH RISKS

Definition
"Knowing about" FGC refers to possessing specific factual information about the procedure, which may or may not affect attitudes toward it. The specific items to be tested may differ from one country to another, but illustrative knowledge items include the following.

Legal: Is the practice of female circumcision legal or illegal in this country?

Religious: Does the Islamic faith require girls to be circumcised?[2]

Health risks: Are girls who undergo FGC at greater health risk than those who do not?

This indicator is calculated as:

$$\frac{\text{\# of respondents that know about the (\textit{legal status/religious position/health risks}) of FGC}}{\text{Total \# of respondents}} \times 100$$

The local term for female circumcision is generally used in this type of question, rather than the more technical WHO classification presented above.

Data Requirements
Response to questions on survey

Data Source(s)
Representative survey of the population

Purpose and Issues
An important first step in eradicating FGC is to raise awareness about the procedure and to expel widely held myths. Two key points that are useful to this end are (1) that FGC is illegal in a given country and (2) that FGC is not mandated by Islam.

A third point often used by advocacy groups relates to the negative health consequences of FGC. Indeed, the number one objection to this practice in studies conducted to date has been negative health consequences. As such, it would seem logical for evaluators to test on this knowledge item. However, the issue of the negative health consequences of FGC is not clear-cut. As indicated in the introduction to this section, the review by Obermeyer (1999) documented short-term effects (pain, hemorrhage, shock, and infection) and long-term effects (urinary infection, scarred tissue, fertility problems, and complications during child birth); however, the evidence from the review suggests that severe complications are relatively infrequent. Obermeyer concludes that negative health consequences may in fact exist, but cautions that the available evidence is less compelling than anecdotal accounts suggest. Thus, if evaluators are testing "accurate knowledge" about FGC, they must be vigilant not to overstate the frequency of negative health consequences, even if anti-FGC advocates cite such consequences as one of the primary reasons to ban the practice.

Also, information on the legal status of FGC may strengthen the resolve of community members to discourage the practice. And information on the position of the Islamic faith on FGC may dispel the widely held myth that women of Islamic faith must be circumcised.

These questions also serve as useful markers of progress in the wake of public information campaigns designed to increase awareness of FGC and to combat misconceptions about the practice.

However, even with useful and accurate knowledge of FGC, people may not change their attitudes or behavior. In areas where strong social convention affecting women's roles keeps FGC in place, those forces may predominate despite improved knowledge.

Evaluators should break down indicators measuring knowledge by age, sex, and education to better understand differences among these subgroups.

---

[2] The answer is no.

PERCENT OF THE POPULATION FAVORABLE TO THE CONTINUATION OF FGC

Definition

"Continuation" refers to perpetuation of the practice of FGC. "Favorable" is operationally defined by items such as those listed below.

This indicator is calculated as:

$$\frac{\text{\# of respondents favorable to the continuation of FGC}}{\text{Total \# of surveyed}} \times 100$$

Data Requirements

Response to questions on survey

Data Source(s)

Large-scale representative survey (e.g., DHS)

Purpose and Issues

The question "do you think that female circumcision should be continued, or should it be discontinued?" (from the DHS module on female circumcision) provides a useful indicator of public acceptance of FGC in a given country. Evaluators can ask it of all respondents, male or female, in a survey, and can present the responses separately for men and women. A similar question, appropriate for women with daughters, is "do you intend to have any of your daughters circumcised?" or "have any of your daughters been circumcised?"

Responses to these items from representative surveys of the population serve two purposes: (1) they indicate the beliefs that public information campaigns must address the elimination of this practice, and (2) they serve as "markers of progress" if data are available over two or more surveys.

Any effort to abolish female circumcision must take into account beliefs that are widely held by members of the target population. The DHS module on female circumcision allows for the measurement of specific beliefs that support the continuation of FGC. Other sources cite the same beliefs in slightly different terms. For evaluation purposes, one can track changes in attitude regarding FGC through population-based surveys. An illustrative set of beliefs favorable to the continuation of FGC includes the following:

- Men prefer women who are circumcised (better marriage prospects);
- Islam/religion requires female circumcision;
- Circumcision is a good tradition/part of our cultural heritage;
- Circumcision is important to avoid the wrath of the ancestors;
- Female circumcision gives greater pleasure to the husband;
- Female circumcision preserves virginity;
- Female circumcision prevents adultery; and
- Circumcision is an important part of gender identity for women.

Gender Implications of this Indicator

FGC is a traditional practice whose basis is to control female sexuality and to make a woman "marriageable." Although FGC violates the human rights of women and girls, women may be as likely as men to support the continuation of the practice as long as they believe the myths perpetuated to support the practice (e.g., that uncircumcised women are "unclean," cannot give birth, and are promiscuous). Many anti-FGC groups are now trying to reach men, women, boys, and girls with information that counteracts the myths with facts, exposes the health risks, and fosters an understanding of gender equity and human rights.

PERCENT OF WOMEN 15-19 YEARS OLD WHO HAVE UNDERGONE FEMALE CIRCUMCISION

Definition
This indicator refers to all forms of FGC described in the introduction. The question in the DHS module reads, "Have you yourself ever been circumcised?"

This indicator is calculated as:

# of women 15-19 who report having undergone female circumcision $\times$ 100
Total # of women 15-19 surveyed

Data Requirements
Self-report; responses to question on survey

Data Source(s)
Large-scale representative survey (e.g., DHS)

Purpose and Issues
This indicator measures the success of programmatic initiatives in reducing the practice of FGC. Although the goal of eradication programs is to eliminate FGC entirely for all age groups, change can most readily be detected by focusing on the 15-19 year old group. In most societies that practice FGC, the procedure is performed before or around puberty; thus, any reduction in the incidence of the practice will first be apparent among this age group.

If change occurs among this age group on this variable, evaluators can further analyze this change by education levels, geographic location, religion, and other variables that may help explain the change and may identify the innovators.

One important caveat relates to the sample used for the survey. In countries where the DHS is limited to married women (e.g., Egypt), drawing conclusions about trends in the practice of FGC may be misleading because those who marry as teenagers are more likely to be circumcised than those who marry later. In populations where FGC is declining, comparisons of DHS data on FGC across age cohorts will fail to show the changes that were actually taking place. In short, the sample must include all women 15-19 in estimating the prevalence of FGC with DHS or other survey data (Mensch, 2001).

A related indicator is the age at circumcision. In Ghana, laws prohibiting the practice may drive it underground, and one outcome may be circumcision at younger ages in life. Evaluation of programs to eradicate FGC should track this variable as a possible unintended consequence.

Part III.M Reproductive Health in Emergency Situations

- Number of incidents of sexual violence reported per 10,000 population
- Percent of health facilities with adequate supplies for universal precautions
- Number of condoms distributed per 1,000 population
- Number of clean delivery kits distributed

REPRODUCTIVE HEALTH IN EMERGENCY SITUATIONS

Historically, humanitarian agencies responding to emergencies – war, civil strife, famine, environmental disasters – did not think about the reproductive health needs of the people they were serving. They focused on providing shelter, food, water, and health care to prevent deaths due to infectious diseases.

Recognition of the importance of reproductive rights and needs of persons affected by armed conflict has, however, evolved rapidly since 1993. Several events highlighted the lack of attention: a 1993 *Lancet* editorial argued for reproductive freedom for refugees; a 1994 survey of refugee settings by the Women's Commission for Refugee Women and Children exposed the virtual lack of reproductive health services; in 1994, Rwandan women refugees in Tanzania and Zaire (now the Democratic Republic of Congo) demanded access to, at least, the services they had used before their flight (*The Lancet*, 1993 & Wulf, 1994). The November 1994 International Conference on Population and Development in Cairo, which specifically articulated refugees' right to reproductive health services, provided a major impetus to this evolution (United Nations, 1995).

In many important respects, refugees are no different from the people in stable settings who have been the focus of development efforts and reproductive health programs for decades. Indeed, before flight from their homes and villages, they may have been the very individuals who participated in such programs. Thus, sound principles of program design, monitoring, and evaluation developed in stable settings may also apply to refugee settings.

However, refugees' experiences of conflict, flight, and displacement introduce factors that program planners must consider as they establish reproductive health services. The poverty, powerlessness, family dissolution, and social instability characteristic of refugees' lives may affect their reproductive health desires, their exposure to risk – of STIs and violence particularly – and their capacity to act (McGinn, 2000). Reproductive health programs must consider the living situations of all those

affected by the conflict, not only the displaced themselves. The displaced may stay in segregated refugee camps or live intermingled with the local population – who are often materially little better off than are the refugees – in villages, towns or cities. In either case, the refugees and the humanitarian organizations that come to serve them change the social, political, economic, and physical environment. Programs must address the issue of equity and the potential for inter-group tension as they determine the services they will offer, and to whom.

The acknowledgment of reproductive rights for refugees places a burden of responsibility upon humanitarian actors to provide the health services refugees need to exercise these rights.

Definition of Terms and Phases of Conflict

The term "refugee," in legal language, refers to a person who has fled his or her home, has crossed an international border, and is unable or unwilling to return because of persecution based on race, religion, nationality, membership in a particular social group, or political opinion. The term also informally includes persons fleeing war, civil strife, famine, and environmental disasters. "Internally displaced persons" (IDPs) have been forced from their homes but remain within the borders of their own countries. Because countries in conflict are often unable or unwilling to provide needed health and social services to IDPs and because the international community may be averse to overstepping the sovereign rights of states, IDPs may receive little international attention, and victims may go unprotected and unassisted.

All those affected by armed conflict are persons of concern: these include refugees, internally displaced persons, and the host populations residing in the locations of asylum. In general, the terms used here – "refugees," "displaced," and "war-affected" – refer to all these affected groups unless otherwise noted.

Phases of Conflict

Complex humanitarian emergencies often fall into phases for guidance in determining program needs and setting priorities. The diagram below (Busza and Lush, 1999) is a useful description of commonly discussed phases. However, conflict is rarely a linear process. A region in conflict often exhibits characteristics of more than one phase at a time as it moves back and forth through the phases.

In the exodus/emergency and post-emergency phases, assessment is typically limited to measuring inputs and functional outputs, specifically logistics. Ironically, evaluators also routinely collect good data on mortality – a long-term outcome measure – at least in closed camp settings, though data are often not age, sex-, or cause-specific. As the population moves into the stabilization and later phases, the data collection system may expand to cover other aspects of the supply or demand environment. In these phases, the programs and monitoring systems may resemble those in development settings and may face many of the same data collection challenges.

Conflict Phase	Description
Exodus / Emergency	The event (such as outbreak of war or escalation of violence) that causes flight, followed by loss of infrastructure, essential services, and the breakdown of political and social organization. At this stage, mortality, morbidity, and malnutrition are at their highest, and people may lack access to supplies for basic needs.
Post-Emergency	National and international aid responses have begun to have impact, and most basic subsistence needs are being met regularly. Some informal infrastructure and routine have been established, although political and social stability may still be precarious.
Stabilization	More services have been restored and people have adjusted to life in displacement. This phase can continue for many years.
Settlement/Repatriation	Although permanent settlement in a host country or repatriation results in very different experiences for the people involved, both signal a degree of resolution to the conflict, and frequently an end to dependence on aid agencies. New lifestyles are forged or old ones re-established, and the needs of the community become ones of long-term development.

Methodological Challenges of Evaluating Reproductive Health Programs in Emergency Situations

- **The destruction of infrastructure and systems in war limits providers' ability to deliver services and evaluate programs.**

By its very nature, armed conflict is destructive. Indeed, a strategy of war is to destroy infrastructure – roads, communication, utilities, health facilities. Agencies working in conflict settings must therefore start from scratch to establish service delivery systems and systems to evaluate the effectiveness of services. Delivery of services receives priority over evaluation.

- **Refugee populations move, and thus measurement becomes difficult.**

Refugees often move more than once. In the early phases of a conflict, large-scale movement can occur in stages as individuals and families make their way to safe havens. Once in a "stable" setting, however, influxes and egresses are common as some family members leave to find work, to return home to harvest their fields, or to test other relocation sites. The ever-shifting denominator complicates measurement.

- **Refugees may view data collection as coercive, and thus data quality may be compromised.**

Although a danger of courtesy bias and intimidation in data collection exist in any population, refugees and the displaced depend for their very lives on the agencies seeking information from them. They may perceive that their participation and their responses will determine access to services that are fundamental to survival.

- **Data collection is limited to accessible populations.**

The published and unpublished literature is biased towards refugees living in stable camp settings, simply because access to other groups – the displaced, those living in scattered sites, those living in insecure areas – is often difficult or impossible. Safety concerns, such as land mines and contact with armed combatants, and practical concerns, such as the inadvisability of traveling at night and the lack of accommodations, limit the ability of staff to travel to deliver services or to collect data. Program staff must guard against generalizing their findings to all refugees and displaced.

- **Agencies may not coordinate monitoring efforts among themsleves.**

The office of the United Nations High Commissioner for Refugees (UNHCR) is the intergovernmental agency responsible for the well-being of refugees (except for internally displaced persons, for whom their own government is responsible). UNHCR, and other coordinating agencies, such as OCHA (UN Office for the Coordination of Humanitarian Affairs), work through many reputable humanitarian and governmental organizations, each of which has its own mission and donors, and finance, personnel, logistics, and record-keeping systems. Efforts to create a common monitoring system or to coordinate data collection at the field level may be made by these coordinating bodies or by the organizations themselves in a particular setting, but the task is complex. International, interagency minimum standards for disaster response, developed through the Sphere Project, have facilitated this task. First articulated in 1998, the standards are voluntarily adopted by humanitarian agencies, and their use is spreading gradually (The Sphere Project, 2000).

- **Technically competent staff are in short supply.**

International response at the onset of emergencies includes well-trained medical staff to provide some services. However, it is always necessary to engage staff or volunteers from among the refugee population for many tasks; these are the majority of workers, especially as the situation stabilizes and the emergency agencies phase out. Frequently, however, refugees with education and technical skill are *not* the ones who remain in refugee camps; their social networks provide them with more attractive and safer alternatives. This exodus applies in particular for trained health workers and persons with research or data analysis experience.

- **Humanitarian agencies and donors plan for the short term.**

Humanitarian agencies are expert in immediate response to emergencies. This is their mission; most are not long-term development agencies. Yet, most of the refugees and displaced in the world are in the stabilization phase, and long-term program objectives are appropriate. Donors and relief agencies typically plan in 6- and 12-month cycles, making measurement of intermediate and long-term outcomes impractical. Ironically, a program is of-

ten funded for 3 or more years, but in 6-month increments, with only the most basic functional outputs measured.

However, several agencies have both an immediate relief as well as a longer-term program function. Many agencies have also worked to ensure that their immediate response is consistent with longer-term program needs to ease the transition in services and data collection and use.

- **We have limited program and research experience on reproductive health in forced migration situations to guide us.**

We are only beginning to understand the effects of forced migration on reproductive health knowledge, attitudes, practice and, ultimately, on reproductive health status. We have limited experience with how to gather information related to movement, mental health, family dissolution, and social change, yet these factors may be key to understanding the needs, desires, resources, and concerns of populations affected by war. Many people debate the ethics of asking questions of traumatized people for whom services may still be unavailable.

We expect that as we gain experience and find answers, this information will be useful to improve services to both refugee and stable populations.

The Indicators

The indicators included in this section are those developed by a community of agencies* – multilateral organizations, governmental agencies, non-governmental organizations, universities – in a guide entitled, *Reproductive Health in Refugee Situations: An Inter-Agency Field Manual* (UNHCR, 1999). The manual recognizes the classic construct of stages of conflict and recommends implementing a **Minimum Initial Service Package** during the emergency phase – i.e., the period of days or weeks at the beginning of a refugee crisis.

The indicators described in detail here are those recommended for monitoring implementation of this **Minimum Initial Service Package** only. The *Field Manual* also recommends that comprehensive reproductive health services be put in place as soon as possible after the emergency phase and includes indicators for monitoring these more extensive services. A summary list of these indicators is included here with reference to parallel indicators elsewhere in this compendium.

---

* The Inter-Agency Working Group on Reproductive Health in Refugee Situations comprises over 30 humanitarian and reproductive health groups who meet regularly to discuss progress and needs in the field. The group was instrumental in developing and revising *Reproductive Health in Refugee Situations: An Inter-Agency Field Manual.*

NUMBER OF INCIDENTS OF SEXUAL VIOLENCE REPORTED PER 10,000 POPULATION

Definition
This indicator is calculated as:

$$\frac{\text{\# of incidents of sexual violence reported in specified reference period}}{\text{Total camp population}} \times 10,000$$

In the emergency phase, sexual violence incidents are most commonly defined as rape.

Data Requirements
Information on the number of incidents of sexual violence reported within a specified period of time (e.g., 6 months) and information on the total number of people in the refugee camp

Data Source(s)
Reports of sexual violence incidents filed with any authority, such as UNHCR protection or other staff, police, local authorities, or health facility staff

Purpose and Issues
The term "sexual violence" covers "all forms of sexual threat, assault, domestic violence, interference and exploitation including involuntary prostitution, statutory rape and molestation without physical harm or penetration" (UNHCR, 1995). In the emergency phase, rape is the form of sexual violence that receives most attention. Note, however, that reproductive health programs should include prevention and response to other forms of sexual violence, as well as gender-based violence, after the emergency phase.

Sexual violence is strongly associated with situations of forced population movement. In this context, all actors in the emergency response must be aware of this issue and preventive measures must be put in place. The UNHCR's *Guidelines for Prevention and Response to Sexual Violence Against Refugees* (1995) should be adhered to in the emergency response. Measures for assisting refugees who have experienced sexual violence, including rape, must also be established in the early phase of an emergency.

Women who have experienced sexual violence should be referred for health services as soon as possible after the incident. Protection staff should also be involved in providing protection and legal support to survivors of sexual violence.

Key actions to reduce the risk of sexual violence and respond to survivors during the emergency, include the following:

- Design and locate refugee camps, in consultation with refugees, to enhance physical security;
- Ensure the presence of female protection, health staff, and interpreters;
- Include the issues of sexual violence in the health coordination meetings;
- Ensure refugees are informed of the availability of services for survivors of sexual violence;
- Provide medical response to survivors of sexual violence, including emergency contraception as appropriate; and
- Identify individual or groups who may be particularly at risk to sexual violence (single female heads of households, unaccompanied minors, among others), and address their protection and assistance needs.

Where possible, the evaluator should obtain data on age and sex-specific incidence rates.

Gender Implications of this Indicator

Reported sexual violence in emergency settings almost always involves female victims and male assailants. Sexual violence against women occurs during all phases of an emergency:

(1) Emergency
Sexual violence has been an instrument to persecute, humiliate, torture, and dominate women and their families. Systematic and politically motivated sexual violence has led many communities and individuals to seek asylum in other areas or countries (UNHCR, 1993).

(2) Exodus
Women are particularly vulnerable during the process of relocation while crossing military lines, areas of civil conflict, and borders. Perpetrators at this stage are most likely to include bandits, smugglers, border guards, police, members of military forces, and civilians from the host population.

(3) Post-Emergency/Stabilization/Settlement
Refugee or internally displaced women have been subjected to many forms of sexual violence: rape, sexual extortion, sexual molestation and threats, and forced prostitution. Women in emergency situations are vulnerable when regarded as sexual property by male refugees and camp guards, when coerced to have sex in return for basic needs, and when subjected to attackers while traveling long distances to ration distribution points (UNHCR, 1999).

Many acts of sexual violence against women in emergency settings go unreported because women fear acts of retribution, are ashamed, fear rejection by a spouse or by the community, feel powerless, lack support, or distrust public/refugee services. A gender appropriate response to sexual violence includes the presence of female medical staff to attend women who have been exposed to sexual violence, a safe environment for reporting sexual violence that respects confidentiality, and integrated care for women exposed to sexual violence (including medical care, psychosocial care, and protection) [UNHCR, 1999].

PERCENT OF HEALTH FACILITIES WITH ADEQUATE SUPPLIES FOR UNIVERSAL PRECAUTIONS

Definition
The number of facilities equipped for universal precautions

This indicator is calculated as:

$$\frac{\text{\# of health facilities with adequate supplies to carry out universal precautions}}{\text{\# of camp service delivery points}} \times 100$$

Each service-delivery point must define adequate supplies based on the number of potential exposures.

Universal precautions refer to the measures (outlined below) to prevent the transmission of HIV.

Data Requirements
Information on number of health facilities within the refugee camp that have adequate supplies for universal precautions, and information on the total number of camp service-delivery points

Data Source(s)
Inventory and commodities report of camp service-delivery points

Purpose and Issues
Those in charge must emphasize universal precautions against the spread of HIV/AIDS within the health care setting during the first meeting of health coordinators of the refugee camp. Under the pressure of an emergency situation, the field staff may be tempted to take short cuts in procedures and thus to jeopardize the safety of patients and staff. Staff must respect universal precautions. This indicator measures the effectiveness of distribution systems for supplies related to universal precautions.

The guiding principle behind universal precautions to prevent transmission of HIV within the health care setting is that one should assume that all blood, blood products, and bodily fluids are potentially infectious.

The minimum requirements for universal precautions are:

- Facilities for frequent hand-washing;
- Availability of gloves for all procedures involving contact with blood and other bodily fluids;
- Availability of protective clothing;
- Safe handling of sharp objects;
- Safe disposal of waste materials;
- Appropriate cleaning, disinfecting, and sterilizing of medical instruments;
- Proper handling of corpses; and
- Appropriate handling of workplace injuries (UNHCR, 1999).

NUMBER OF CONDOMS DISTRIBUTED PER 1,000 POPULATION

Definition
The volume of condoms distributed in relation to the population of the camp

This indicator is calculated as:

$$\frac{\text{\# of condoms distributed in a reference period}}{\text{Total population of the camp}} \times 1000$$

Data Requirements
Total number of condoms distributed in the refugee camp within a specified period of time (i.e., one month); the total population in the refugee camp

Data Source(s)
Condom distribution lists

Purpose and Issues
Availability of condoms should be ensured from the beginning of the emergency so that they can be provided to anyone who requests them. Sufficient supplies should be ordered to cover potential need.

As well as providing condoms on request, field staff should make sure that refugees are aware that condoms are available and where they can be obtained. Condoms should be made available in health facilities, especially when treating cases of STIs. Other distribution points should be established so that those requesting condoms can obtain them in privacy.

One limitation of this indicator is that distribution does not necessarily equate to use, especially where the product is given away free of charge. A second caveat is that in refugee populations with a high proportion of children, the number of condoms per 1,000 would decrease, making comparisons across refugee settings invalid. Nonetheless, as a crude measure of protection against unwanted pregnancy and disease prevention, this indicator is potentially useful, especially since data are readily available.

NUMBER OF CLEAN DELIVERY KITS DISTRIBUTED

Definition
This indicator is calculated as:

$$\frac{\text{\# of clean delivery kits distributed}}{\begin{array}{l}\text{Estimated \# of pregnant women}\\\text{in the refugee camp}\end{array}} \times 100$$

Delivery kits can be those for use by mothers or birth attendants as well as those for use by midwives.

Data Requirements
Information on the total number of clean delivery kits distributed within the refugee camp and information on the estimated number of pregnant women within the camp (estimated to be 75-125 in a three-month period in a population of 10,000)

Data Source(s)
Distribution or inventory lists of delivery kits and health facility data on the number of pregnant women within the camp

Purpose and Issues
A refugee population will include women in the later stages of pregnancy who will deliver within the initial phase of the emergency. Camp personnel should provide simple delivery kits for home use to women in the late stages of pregnancy. Women themselves or traditional birth attendants (TBAs) can use these very simple kits. Staff can assemble kits on site, which should include: one sheet of plastic, two pieces of string, one clean razor blade, one bar of soap, and a cotton cloth to prevent hypothermia in the newborn.

Evaluators can use a formula based upon the crude birth rate (CBR) to calculate the supplies and services required. With a crude birth rate of 3 to 5 percent per year, some 75-125 births will likely occur in a 3-month period in a population of 10,000. Using this estimate, staff can calculate the number of kits they should order.

In the early phases of an emergency, births will often take place outside the health facility without the assistance of trained health personnel. Approximately 15 percent of births will involve some complications. Those assisting in the birth should refer complicated births to the health center. The supplementary unit of the New Emergency Health Kit (NEHK-98) has all the materials needed to ensure safe and clean normal deliveries. The health center can manage many obstetric emergencies with the equipment, supplies, and drugs contained in the NEHK-98. When the centers cannot manage obstetric complications, they should stabilize the patients before transferring them to the referral hospital.

Reproductive Health in Refugee Situations Indicators and
Compendium of Reproductive Health Indicators

Note: Many of the indicators for the post-emergency and later phases of refugee crises correspond to the areas of reproductive health outlined in this *Compendium*. The full list of indicators for these later stages included in *Reproductive Health in Refugee Situations, An Inter-Agency Field Manual,* is presented in the first column. Each one is then cross-referenced with the indicators described elsewhere in the *Compendium*. In some cases, they are exactly the same; in others, they are similar. Column 3 gives the section in which each indicator appears in this manual.

Indicators in Reproductive Health in Refugee Situations, An Inter-Agency Field Manual	Parallel Indicator in this Compendium	Section in this Compendium
Newborn Health		
1. Neonatal mortality rate	Neonatal mortality rate (NMR)	III.E
2. Percent of births which are of low birth weight (<2500 g)	Percent of live births with low birth weight	III.E
3. Percent of births which are of very low birth weight (<1500 g)	Percent of live births with low birth weight	III.E
4. Stillbirth ratio		
Safe Motherhood		
1. Percent of reported maternal deaths investigated according to established guidelines and results of which are disseminated to health staff	Percent of facilities that conduct case review/audits into maternal death/near misses	III.D
2. Percent of women attending antenatal services at least once	Percent of women attended at least once during pregnancy for reasons related to the pregnancy	III.D
3. Percent of women delivering tested for syphilis during pregnancy	Percent of pregnant women attending antenatal clinics screened for syphilis	III.D
4. Percent of pregnant women screened for syphilis who test positive	Percent of pregnant women attending antenatal clinics screened for syphilis	III.D
5. Incidence of unsafe and spontaneous abortions	Abortion rate (AR) and total abortion rate (TAR)	III.I
6. Percent of women delivering adequately vaccinated with tetanus toxoid	Percent of pregnant women with at least two doses of tetanus toxoid immunization	III.E
7. Percent of women with obstetric complications		
8. Percent of women with obstetric emergencies treated in timely and appropriate manner	Percent of women with obstetrical complications treated within two hours at a health facility	III.D
9. Percent of women delivering attended by trained health worker	Percent of births attended by skilled health personnel	III.D
10. Percent of WRA who can name at least 2 danger signs of obstetric complications	Percent of audience that know three primary warning/danger signs of obstetric complications	III.D
11. Percent of women delivering by Caesarean section	Cesarean sections as a percent of all live births	III.D
12. Percent of women with abortion complications treated in timely and appropriate manner		
13. Percent of women delivering who receive at least 1 postpartum care visit	Percent of women attended during the postpartum period by skilled personnel	III.D
14. Percent of newborns receiving BCG and polio by 1-month birthday		

Indicators in Reproductive Health in Refugee Situations, An Inter-Agency Field Manual	Parallel Indicator in this Compendium	Section in this Compendium
Sexual Violence (SV)		
1. Percent of SV survivors who receive basic set of psychosocial and medical services	Number of VAW service visits provided, by type of service, in a reference period	III.K
2. Percent of SV survivors who present for care within 3 days of event		
3. Percent of identified SV offenders prosecuted to the full extent of the law		
4. Percent of designated health workers trained within past 2 years to provide services to SV survivors	Number of trainees by type of personnel and topic of training	II.D
STI/HIV/AIDS		
1. Percent of blood samples for transfusion tested for HIV	Percent of transfused blood units screened	III.C
2. Incidence of STIs		
3. Percent of STI patients assessed and treated according to protocol	Percent of STI patients appropriately diagnosed and treated	III.C
4. Percent of designated health workers trained to manage STI cases according to protocol	Number of trainees by type of personnel and topic of training	II.D
5. Percent of health workers who demonstrate use of universal precautions		
6. Percent of potential condom outlets with condoms available	Condoms available for distribution nationwide	III.C
7. Percent of persons in target population who recognize a condom, know its preventive effects, and can describe how to use it properly	Percent of population who know HIV prevention methods	III.C
8. Percent of persons in target population reporting using condom at last intercourse with non-regular partner	• Condom use at last higher risk sex • Young people using a condom at last higher risk sex • Percent of sexually initiated adolescents who used a condom at first/last sex • Percent of sexually active, unmarried adolescents who consistently use condoms	III.C III.C III.H III.H
Family Planning		
1. Contraceptive prevalence rate (CPR)	Contraceptive prevalence rate (CPR)	III.B
2. Percent of health workers who provide family planning services trained in the past 2 years	Number of trainees by type of personnel and topic of training	II.D
3. Percent of sexually active refugees able to cite major messages about family planning	Percent of audience who recall hearing or seeing a specific message	II.F
4. Percent of service-delivery points maintaining a minimum of 3-months supply of each of combined oral contraceptive pills, progestin-only pills, and injectables	Percent of facilities whose stock levels ensure near-term product availability	II.E

Indicators in Reproductive Health in Refugee Situations, An Inter-Agency Field Manual	Parallel Indicator in this Compendium	Section in this Compendium
RH of Young People		
1. Incidence of STIs in young people	• Percent of adolescents who were ever diagnosed with an STI	III.H
2. Percent of births to young women		
3. Contraceptive prevalence rate among young people	Contraceptive prevalence rate (CPR)	III.B
4. Percent of sexually active young people reporting condom use at last intercourse	• Percent of sexually initiated adolescents who used a condom at first/last sex • Percent of sexually active, unmarried adolescents who consistently use condoms	III.H III.H
5. Percent of young people assessed, treated, and counseled according to protocol		

CONTRIBUTORS TO THE *COMPENDIUM OF INDICATORS*[1]

Subject Area	Contributors	Organization
Part I – Overview of the *Compendium*	Ties Boerma Richard Greene Sara Pacqué-Margolis Amy Tsui Norma Wilson	MEASURE *Evaluation*/UNC Population Leadership Program USAID Gates Institute/John Hopkins University USAID/Office of Population
Women's Status and Empowerment	Michal Avni Suzanna Stout Banwell Mai-Britt Dohlie Mary Ellsberg Susheela Engelbrecht Alfredo Fort Victoria Frye Sreela das Gupta Meg Greene Karen Hardee Judith Helzner Britt Herstad Jodi Jacobson Sunita Kishor Kathleen Kurz James McMahan Sarah Martin Constance Newman Karin Ringheim Diana Santillán Sid Schuler Yvonne Sidhom Nancy Yinger	USAID CHANGE PRIME II/Intrah Women's Health Exchange PRIME II/Intrah PRIME II/Intrah CHANGE Tulane University PAI POLICY/The Futures Group IPPF/WHR MEASURE Communication/PRB CHANGE MEASURE DHS+/Macro International Reproductive Health and Nutrition Team/ICRW PRIME II/Intrah MEASURE *Evaluation*/JSI PRIME II/Intrah Women's Reproductive Health Initiative/PATH Empowerment of Women Research Program/JSI Empowerment of Women Research Program/JSI PRIME II/Intrah MEASURE Communication/PRB
The Policy Environment	Harry Cross Karen Foreit Karen Hardee Jeff Jordan Nancy McGirr John Stover	POLICY/The Futures Group POLICY/The Futures Group POLICY/The Futures Group POLICY/The Futures Group POLICY/The Futures Group POLICY/The Futures Group

[1] The persons on this list provided some input into the respective sections, ranging from drafting and finalizing the text to commenting on an isolated indicator in the section. The appearance of their names should not be interpreted as their endorsement of all ideas in the relevant section.

Management	Jaime Benavente	Center for Health Services/MSH
	Joseph Dwyer	Management and Leadership Prog./MSH
	Sarah Johnson	Management and Leadership Prog./MSH
	Scott McKeown	INFORM Program/MSH
	Marc Mitchell	Harvard School of Public Health
	Mary O'Neil	Management and Leadership Prog./MSH
	Anne Young Packham	Management and Leadership Prog./MSH
	Linde Rachel	Management and Leadership Prog./MSH
	Gerry Rosenthal	Center for Health Reform and Finance/MSH
	Steve Sapirie	Inform Program/MSH
	Joseph Seale	Private management consultant
	Judith B. Seltzer	Center for Health Reform and Finance/MSH
	Mary Taylor	Independent Consultant
	Pamela Teichman	Pathfinder
	Barbara Tobin	Management and Leadership Prog./MSH
Training	Thada Bornstein	QAP/URC
	Sue Brechin	JHPIEGO
	Barbara Felling	DELIVER/JSI
	Alfredo Fort	PRIME II/Intrah
	Kama Garrison	JHPIEGO
	Linda Ippolito	Intrah
	Cathy Murphy	PRIME II/Intrah
	Jana Ntumba	MSH
	Robert Rice	FHI
	Sharon Rudy	PRIME II/Intrah
	Lois Schaefer	JHPIEGO
	Cathy Solter	Pathfinder
	Rick Sullivan	JHPIEGO
	Tim Williams	DELIVER/JSI
Commodities and Logistics	Dana Gelfeld Aronovich	DELIVER/JSI
	Tim Johnson	Division of Reproductive Health/CDC
	Steve Kinzett	DELIVER/JSI
	Rich Owens	DELIVER/JSI
	Sandhya Rao	DELIVER/JSI
	Lea Teclemariam	DELIVER/JSI
	Tim Williams	DELIVER/JSI
Behavior Change Communication	Marcela Aguilar	PCS/JHU/CCP
	Stella Babalola	PCS/JHU/CCP
	Marc Boulay	PCS/JHU/CCP
	Maria Elena Figueroa	PCS/JHU/CCP
	Nathalie Likhite	MSH
	D. Lawrence Kincaid	PCS/JHU/CCP
	Sarah Landon	PCS/JHU/CCP
	Theresa Norton	Learning and performance support office/ JHPIEGO Corporation

Behavior Change Communication (cont.)	Suruchi Sood Carol Underwood Kim Witte E. Kevin Zembower	PCS/JHU/CCP PCS/JHU/CCP PCS/JHU/CCP PCS/JHU/CCP
Operations Research	Sarah Harbison Marge Horn Celeste Marin John Townsend	USAID USAID FRONTIERS/Tulane University FRONTIERS/Population Council
Service Delivery Environment – Quality of Care	Karen Beatty Nancy Fronczak Gary Lewis Mandy Rose Jim Shelton Krista Stewart Tara Sullivan Ann Way Tim Williams	EngenderHealth MEASURE DHS+/Macro International PCS/JHU/CCP MEASURE *Evaluation*/Macro International USAID USAID MEASURE *Evaluation*/Tulane MEASURE DHS+/Macro International DELIVER/JSI
The Quality Assurance Approach (Applied to Maternal Health)	Diana Silimperi Bruno Bouchet Stephane Legros	QAP/URC QAP/URC QAP/URC
Gender Equity/ Sensitivity	Mai-Britt Dohlie Susheela Engelbrecht Alfredo Fort Meg Greene Karen Hardee Sunita Kishor James McMahan Constance Newman Karin Ringheim Yvonne Sidhom	PRIME/Intrah PRIME/Intrah PRIME/Intrah PAI POLICY/Futures Group MEASURE DHS+/Macro International PRIME/Intrah PRIME/Intrah Women's Reproductive Health Initiative PRIME/Intrah
Integration of Services	Ian Askew	FRONTIERS/Population Council
Global Indicators for Reproductive Health	Wilma Doedens Iqbal Shah	WHO/Geneva WHO/Geneva

Family Planning Text for this section was drawn primarily from Bertrand, Magnani and Knowles, 1994.	Silvia Alayon Sîan Curtis Anrudh Jain Robert Magnani Leo Morris John Stover Paul Stupp	MEASURE *Evaluation*/Macro International MEASURE *Evaluation*/Macro International Population Council MEASURE *Evaluation*/Tulane Division of Reproductive Health/CDC The Futures Group Division of Reproductive Health/CDC
STI/HIV/AIDS Text for this section was drawn primarily from UNAIDS, 2000.	Ties Boerma Lisanne Brown John Stover	MEASURE *Evaluation*/UNC MEASURE *Evaluation*/Tulane The Futures Group
Safe Motherhood and Newborn Health	Koki Agarwal Deb Armbuster Patsy Bailey Ties Boerma Pierre Buekens Sian Curtis Wilma Doedens Donna Espeut Joy Fishel Stephen Gloyd Wendy Graham Rik Guidoitti Marge Koblinsky Nazo Kureshy Joy Lawn Jerker Liljestrand Deborah Maine David Marsh Allisyn Moran Nancy Nachbar Monica Parise Elisabeth Ransom Carine Ronsmans John Ross Hazel Simpson Cindy Stanton Anne Tinker Tessa Wardlaw Kim Wheeler Jelka Zupan	The Futures Group Consultant FHI MEASURE *Evaluation*/UNC University of North Carolina MEASURE *Evaluation*/Macro International WHO/Geneva CSTS/Macro International MNH/JHPIEGO University of Washington Dugald Baird Centre/U. of Aberdeen WHO/Geneva NGO Networks NGO Networks International Perinatal Care Unit, Centre for International Child Health, Institute of Child Health World Bank AMDD Program, Columbia University Save the Children Federation/US MNH/JHPIEGO Center for applied behavioral and evaluation research/AED CDC PRB London Sch. of Trop. Med. & Hygiene The Futures Group DFID/Nepal Safe Motherhood Program MNH/JHPIEGO Save the Children UNICEF Western Pacific Regional Office/WHO WHO/Geneva

Women's Nutrition	Eunyong Chung	USAID
	Mary Cogswell	Maternal and Child Nutrition Branch, Division of Nutrition and Physical Activity, CDC
	Leslie Elder	The World Bank, HDNHE
	Nadra Franklin	Linkages/AED
	Larry M. Grummer-Strawn	Maternal and Child Nutrition Branch, Division of Nutrition and Physical Activity, CDC
	Phil Harvey	MOST/Johns Hopkins University
	Sam Khan	USAID
	Jacky Knowles	Maternal and Child Nutrition Branch, Division of Nutrition and Physical Activity, CDC
	Kathy Krasovec	PATH
	Kathleen Kurz	Reproductive Health and Nutrition Team/ICRW
	Miriam Labbok	UNICEF
	Chessa Lutter	PAHO
	John Mason	Tulane University
	Ellen Mathys	Tulane University
	Altrena Mukuria	MEASURE DHS+/Macro International
	Nancy Mock	Tulane University
	Penny Nestel	USAID Micronutrient Global Leadership Project, ILSI
	Abe Parvanta	Division of Nutrition and Physical Activity, CDC
	Ellen Piwoz	SARA/AED
	Sandra Remancus	FANTA Project/AED
Breastfeeding	Nadra Franklin	AED
Adolescent Reproductive Health Programs	Robert Magnani	FOCUS/Tulane
	Kristin Nelson	FOCUS/Tulane
Postabortion Care	Janie Benson	Ipas
	Barbara Crane	Ipas
	Tamara Fetters	Ipas
	Norman Goco	Ipas
	Joan Healy	Ipas
	Charlotte Hord	Ipas
	Ronnie Johnson	Ipas
Male Involvement in Reproductive Health Programs	Sam Clark	PATH
	Adrienne Cox	PHNIP/JSI
	Erin Eckert	MEASURE *Evaluation*/JSI
	Meg Greene	PAI
	Sarah Martin	MEASURE *Evaluation*/JSI
	Alec Moore	MEASURE *Evaluation*/JSI
	Diana Santillán	Empowerment of Women Research Program/JSI
	Nancy Yinger	MEASURE *Communication*/PRB

Violence Against Women	Suzanna Banwell Mai-Britt Dohlie Mary Ellsberg Susheela Engelbrecht Alfredo Fort Victoria Frye James McMahan Constance Newman Yvonne Sidhom Jeanne Ward	CHANGE PRIME II/Intrah PATH PRIME II/Intrah PRIME II/Intrah CHANGE PRIME II/Intrah PRIME II/Intrah PRIME II/Intrah Sexual and Gender-based Violence Initiative, Reproductive Health for Refugees Consortium
Female Genital Cutting	James Phillips Barbara Ibrahim Barbara Mensch Gilda Sedgh Nahla Abdel-Tawab Letitia Reason Patricia Akweongo	Population Council Population Council Population Council Harvard University Independent Consultant Washington University Navrongo Health Research Centre
Reproductive Health in Emergency Situations	Therese McGinn Susan Purdin	Columbia University Columbia University
Performance Improvement	Sue Brechin Alfredo Fort	JHPIEGO PRIME II/Intrah

One criticism of indicator handbooks is that they do not provide sufficient guidance on the selection of indicators. Those new to the evaluation process may wonder how to select from the many different indicators presented in this *Compendium*. To address this issue, we provide a "how-to" guide, intended to show the process in its most basic form. The four main steps in selecting indicators are as follows:

1. State (or formulate) the objectives of the program;

Example: increase percent of sexually active males 15-19 years old who "consistently" use condoms in city "x" in a two-year period.

2. Review the activities to be carried out in pursuit of the objective(s);

3. Develop a simple framework to show how the program will work; that is, how the activities will lead to the desired objective(s) (see framework below).
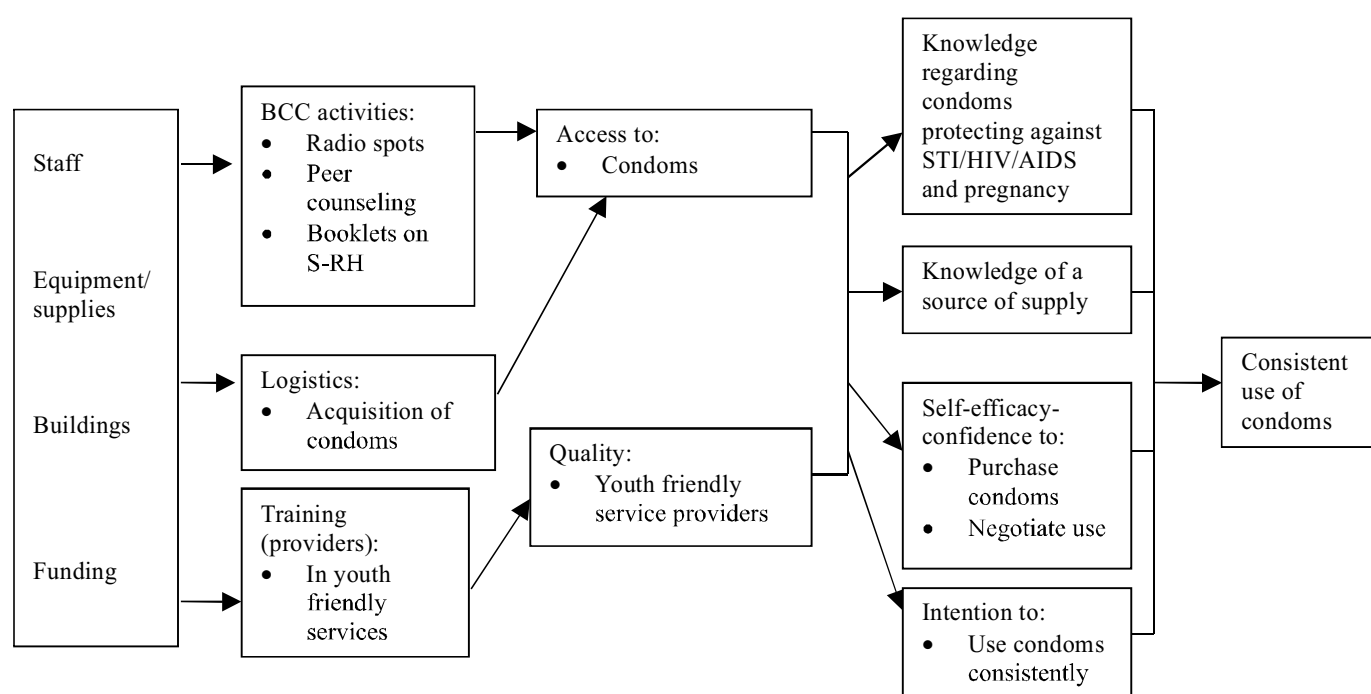
4. Select indicators that measure progress for each (or some) of the boxes.

Example:

Process/functional outputs:[1]
- Number of radio spots produced
- Number of diffusions of radio spots per month
- Number of booklets produced/distributed
- Number of providers trained
- Number of peer educators trained
- Number of peer educators active in program

Service outputs (access and quality):
- Number of outlets that sell/provide condoms to youth in catchment area
- Number of youth friendly facilities in catchment area that counsel youth on sexual and reproductive health
- Quality of service for youth (specific measures will need to be developed)

Results[2]
- Percent of sexually active males 15-19 who:
  - o Know that condoms protect against pregnancy and against STI/HIV/AIDS
  - o Know a source of supply for condoms
  - o Feel confident they could purchase a condom
  - o Feel confident they could negotiate condom use
  - o Intend to use condoms for every sexual relation
  - o Report they always use a condom for sexual relations

Other considerations in selecting indicators:
- Are the data needed to measure the indicators available? If not, are they feasible to collect? (Alternatively, what indicators can be measured with available data?)[3]
- What is the time frame for the evaluation? How often will the program report on the different results? Will the data be available by donor or by program-imposed deadlines?
- What financial support is available for evaluation? Does the organization have funds to conduct a survey? Or does the budget dictate the use of existing data such as service statistics?
- What are the requirements of the donor agency (if applicable)?

For a more detailed discussion of indicators (e.g., how to use multiple indicators, what to measure, how to identify data sources), see Bertrand, Magnani, and Rutenberg, 1996, pages 29-36.

---

[1] Note: Although some organizations track the number of activities completed (i.e., functional outputs), those with greater evaluation capacity focus on results instead. Thus, the BCC section of this *Compendium* (Section II.F) omits "counts" of activities completed.

[2] These results would be considered "outputs" if measured at the program level (among participants/clients in the program) or "outcomes" if measured at the level of the population (among the representative sample youth 15-19 in the general public).

[3] This question represents the "last resort." Whereas the evaluator ideally identifies what a program should measure, he/she may have to settle for what a program can measure and/or use data that has already been collected.

Appendix C
Performance Improvement (PI)

- Indicators for performance improvement

The indicators in Part II.H on the Service Delivery Environment have provided a means of operationally defining access, quality of care, integration of services, and gender equity/sensitivity. This appendix on Performance Improvement (PI) describes a process that can be used in the service delivery context to bring about improvements.

In the past, when organizations experienced performance problems, the common response was to call for more training. Yet even after staff was trained, many of the problems persisted. It became evident that other factors in the work environment influenced performance, and until they were addressed, these problems were likely to continue. In recent years Performance Improvement (PI) has been introduced into the service delivery context as a strategy that takes a systemic, holistic approach to ensuring good performance.

Specifically, Performance Improvement aims at improving service delivery through a multi-step process:

- Identifying gaps between desired and actual performance;
- Analyzing the reasons for these gaps;
- Identifying effective interventions that address these gaps;
- Implementing the interventions; and
- Monitoring performance for improvement in these areas (to determine the extent to which the gap has been closed).

Experience shows that staff needs a certain set of basic inputs to perform well: clear job expectations, immediate performance feedback, adequate physical environment and tools, motivation to do well, support from the organization, and appropriate knowledge and skills (PRIME, 2000).

Although the process will differ from setting to setting, it generally begins when an organization wishes to address its performance problems and decides to enlist the PI process. A Performance Improvement facilita-

tor, experienced with this process, leads the group through a series of clearly defined stages. In the first stage the PI leader, the organization in question, and other stakeholders meet to define the desired outcomes of the process and to discuss the approach itself (who will be involved, in what ways, what will occur, how the activity fits into the goals of the organization). For the PI process to succeed, it is essential to involve relevant stakeholders from the start, including government officials, NGO administrators, providers, clients, and the community. This first stage is important in setting the tone for the project (McCaffery et al., 1999).

In Stage 2, the stakeholders define the desired performance, describe the current performance, and identify the performance gap between the two. The group then prioritizes the performance gaps, completes a root-cause analysis to determine what causes the performance gap, and begins identifying possible interventions to each cause. The PI leader encourages the team to focus on interventions that are practical, sustainable, and cost-efficient.

In Stage 3, the team decides which performance gaps it will address and proceeds to design appropriate interventions. In Stage 4, the PI leader works with the team to ensure that the organization is ready to implement the intervention. Once implementation takes place, the team monitors the process of performance change. In the fifth and final stage, the team evaluates the process, by analyzing whether the performance gap has narrowed during the course of or as a result of the exercise. Ideally, the evaluation method should be designed so that it can be easily integrated into the workplace routine and remain there as a feedback device for workers and managers (McCaffery et al., 1999).

This approach has been tested in a number of countries to date: Burkina Faso, the Dominican Republic, El Salvador, Ghana, Honduras, India, Kenya, Tanzania, Yemen, and others. In each case, a performance problem had been identified, and the PI approach has been used to resolve it.

Methodological Challenges of Evaluating Performance Improvement

- **The evaluation of PI is often done by the same persons involved in the intervention itself, rather than by an external evaluator.**

Because one of the objectives of the exercise is to develop a sense of ownership of the process of improving access and quality, key stakeholders need to be involved at all stages of the PI process, including its evaluation. However, this involvement may color their judgment on the amount of change achieved on different indicators. Two means of addressing this problem are (1) to develop as precise measures of performance as possible, and (2) to involve an external consultant in the evaluation.

- **A number of the indicators are subjective, requiring a judgment on the part of evaluators.**

Some PI indicators are very concrete; for example, percent of providers/staff with an updated written job description for their position. Others – including some that seem measurable – require a subjective judgment. The indicator existence of a staff-oriented leadership of the organization/institution is a case in point. Although the evaluator will gather information from interviews with staff and document reviews, the ultimate "yes/no" for this indicator remains subjective.

Some have argued that in quality improvement initiatives, efforts to ensure local ownership of the PI process are often at odds with methodologically sound data collection procedures and objective measurement (Stinson, 1995). The latter requires "external experts" who are often perceived as outsiders, "inspecting" or "judging" performance in ways that are counterproductive to team-building.

We are not suggesting that the PI process would be better served by limiting the evaluation to more concrete measurements. In fact, the purpose of PI is to influence complex processes that do not lend themselves to easy measurement. At the same time, those implementing PI must be aware of the subjective nature of certain indicators in this type of evaluation.

- **Factors other than PI interventions may affect performance.**

As with other topics in this *Compendium*, attribution is problematic. Other interventions or "naturally occurring" events may affect providers' attitudes or clients' perceptions. Controlling for such confounding factors is difficult at best.

The indicators listed in Table C.1 are organized according to the different inputs required for providers to perform well. Each category is further broken down by indicators applicable at the individual level (of the service provider) and those applicable to the organization in question. Often an indicator at one level will have an analogue at the other.

We should stress that PI is a process. Some would argue that it should not be "reduced" to a set of indicators. "Desired performance" is the end measure of success in the PI process. In contrast to some types of quality improvement tools or audits that measure whether the organization has a strong organizational development system, the PI process focuses more on how organizations achieve success. "Evaluation" in this context serves to document the process as a means of learning from it.

INDICATORS FOR PERFORMANCE IMPROVEMENT (PI)

Definition

No standard list of indicators exists for measuring Performance Improvement; rather, the participating staff of an organization develops the indicators to apply to each specific situation. However, illustrative indicators are listed in Table C.1.

Data Requirements

The indicators generally fall in one of six categories:

- Job expectations;
- Feedback to providers on their job performance;
- Motivation/incentives;
- Job environment;
- Organizational support for providers; and
- Knowledge and skills of providers.

Data Source(s)

Review of documents; interviews with staff; test of knowledge and skills; observation of providers; service and other statistics; program records on absenteeism rates, staff turnover, and utilization of staff time; and special studies

Purpose and Issues

Evaluation is the fifth stage of the PI process. It allows the organization to determine whether it has successfully "narrowed the performance gaps" that it has itself identified during the PI initiative. Methodological challenges are discussed in the preceding text that introduced this topic.

**Table C.1 Performance Improvement (PI) Indicators
(Illustrative for Family Planning)**

I – Indicators of Performance

INDIVIDUAL LEVEL

- Percent of providers/staff performing to standards(e.g., adherence to client provider interaction (CPI) norms when counseling for FP and infection prevention protocols for FP services; making the prescribed number of supportive supervision visits in a quarter)

INSTITUTIONAL LEVEL

- Institution has defined intervention(s) based on a Performance Needs Assessment (PNA)
- Institution with a system to produce and implement a PI plan for its providers/staff
- Institution with logistics/training/supervision/financial/information systems integrated to a comprehensive PI plan for its staff

II – Indicators of the Presence or Absence of the Performance Factors
1. Expectations

INDIVIDUAL LEVEL

- Percent of providers/staff with an updated written job description for their position
- Percent of providers/staff guided by current consensus-derived annual performance objectives
- Percent of providers/staff who can state what is expected of their jobs
- Updated guidelines and protocols are distributed to providers in the areas in which they work

INSTITUTIONAL LEVEL

- Institution systematically uses job descriptions and annual performance objectives to enhance staff performance
- Supervisors review performance expectations on an annual basis and update them as needed

2. Feedback

INDIVIDUAL LEVEL

- Percent of providers/staff who acknowledge receiving feedback (positive and negative) on their performance in the last six months
- Percent of providers/staff who have improved a work habit/procedure based on feedback received in the last year

INSTITUTIONAL LEVEL

- Institution has a feedback system in place to assure providers receive performance feedback
- Institution systematically requires supervisors to provide regular feedback (at least once every six months) to their staff

3. Motivation/incentives

INDIVIDUAL LEVEL
- Percent of providers who state they received positive recognition during the previous year for a specified performance

INSTITUTIONAL LEVEL
- Organization has a functioning system of motivation/incentives for its staff
- Percent of staff that have positive job satisfaction
- Score on individual's job satisfaction rating

4. Job Environment

INDIVIDUAL LEVEL
- Percent of providers/staff who state they have the equipment/tools needed to perform well in their job
- Percent of providers/staff who have adequate tools (light, space, client privacy, audio-visual materials, vehicles, furniture, supplies) to perform the duties expected of them

INSTITUTIONAL LEVEL
- Organization conducts regular assessments to ensure appropriate equipment/tools are available at staff's workplace
- Organization has a system to (re)supply needed materials for providers

5. Organizational Support

INDIVIDUAL LEVEL
- Percent of providers and other relevant staff who received a supportive supervision visit in the last six months
- Percent of providers and other relevant staff strongly identified with their institution
- Percent of providers/staff who are involved in management reviews/decision-making of their clinics/facilities

INSTITUTIONAL LEVEL
- Existence of a staff-oriented leadership of the organization/institution
- Existence of a system that supports providers in the six performance factor areas

6. Knowledge and Skills

INDIVIDUAL LEVEL
- Provider demonstrates specific knowledge and skills to better perform their jobs

INSTITUTIONAL LEVEL
- Existence of a training plan linked to quality assurance and performance improvement plans

LOGISTICS INDICATOR ASSESSMENT TOOL (LIAT)[1]

INTERVIEWER'S GUIDE

Questions #1-4   Note the date, your name and the names of other interviewers, the time you begin the interview and the location of the facility.

Questions #5-8   Circle the type of facility where the survey is conducted and write in the name of the facility.

Question #9   Record the total number of staff at the facility. For warehouses, include staff of the warehouse only, even if it is linked to an SDP such as a district hospital. For SDPs include all staff in the facility, not only those working on logistics.

Question #10   List the titles of all the members of the staff interviewed at the site, how long they held the position and if they have received any training in logistics. Ask about the type of logistics training received by the staff in each product category (e.g., if speaking to the person in charge of immunization ask specifically if he/she has received formal training in cold chain and vaccine logistics).

Questions #11-29   Be sure to read each question exactly as written. If the respondent does not understand the question, you may explain it in your own words or translate to local dialect. If the answer given is not clear or does not clearly fit any response category, use the comment section to provide additional useful information.

Question #30   List all the HIV/test kits by type and brand. Please note if not applicable.

Tables #31-34   Should be filled for each product, following the instructions preceding each table in the questionnaire.

Table 35   Should be filled for each storage area housing products being assessed, following the instructions preceding the table.

Questions #36-40   Answer these questions only if the study team is looking at a cold chain logistics system.

Questions #41-44   Be sure to answer each question exactly as written. If the respondent does not understand the question, you may explain it in your own words or translate to local dialect. If the answer given is not clear or does not clearly fit any response category, use the comments section to provide additional useful information.

Questions #45-46   Record the ending time of the interview and calculate the total time of the interview using the beginning time recorded at the beginning of the questionnaire.

---

[1] DELIVER/JSI, 2002.

FOR SERVICE DELIVERY POINTS, DISTRICT/REGIONAL FACILITIES AND CENTRAL WAREHOUSES

Introduce yourself and all members of the team, including titles/positions. Present the objectives of this assessment and how this interview will help the team to achieve the objectives.

State your objectives here

Explain how the team will conduct the interview, invite relevant interviewees to join the group and begin.

Beginning time of Interview _____	End time of Interview _____

---

1. Date _____ 2. Interviewer(s) _____
3. Region/Province _____ 4**.** District _____
5. Type of facility: (Circle all that apply) a) Urban/Rural b) Public/Private c) Warehouse/Service Delivery Point **(if warehouse go to question 6, if SDP go to question 7)**
6. If warehouse: a) Central b) Regional/Provincial c) District/Zonal
7. If SDP: a)   Hospital b)   Clinic c)   Dispensary d)   Health post e)   Health worker
8. Name of the facility: _____
9. Total number of staff at facility: _____

10. Respondents interviewed at this site:

Title	Length in current position	**Received training in logistics (specify product categories & dates)**

a) _____   _____years/months   _____

b) _____   _____years/months   _____

c) _____   _____years/months   _____

d) _____   _____years/months   _____

e) _____   _____years/months   _____

f) _____   _____years/months   _____

g) _____   _____years/months   _____

*** Logistics includes the following functions: Ordering, Receiving supplies, Inventory management, and supervision. If speaking to the person in charge of EPI, ask specifically "Have you received formal training for cold chain and vaccine logistics?"*

11. Do you use the following logistics forms to manage health products?	
a) Stock cards/records	*If no, go to question 16*
b) LMIS reports (including requisition and issue vouchers, consumption, transactions)	

Please circle answer(s)	Comments

12. How is the information on these forms used? (Circle all that apply)
   a) Calculating consumption
   b) Calculating needs
   c) Reporting use to the higher level
   d) Requesting supplies from the higher level
   e) Other, please explain in comments section	

13. If LMIS reports are used, are they sent to the higher level? a) Yes b) No *(go to question 15)* c) Don't know *(go to question16)* d) Not applicable *(go to question16)*	

Please circle answer(s)	Comments
14. How often are these forms sent to the higher level? a) Monthly b) Quarterly c) Semi-annually d) Annually e) Other (please specify in the comments section) f) Not Applicable	
15. How often are you supposed to send these forms to the higher level? a) Monthly b) Quarterly c) Semi-annually d) Annually e) Other (please specify in the comments section) f) Not Applicable	
16. How many facilities should send reports to this facility? _____ **(if zero, go to Q18)**	
17. Provide an approximate number of facilities that send these reports according to schedule. _____	
18. How many times have you placed an order or submitted a procurement request in the last year? a) None b) 0-3 times c) 3-6 times d) more than 6 times a year (specify reasons in the comments section)	
19. How often are you supposed to place orders or submit a procurement request? a) Monthly b) Quarterly c) Semi-annually d) Annually e) Other (specify in the comments section) f) Not Applicable	
20. Who determines this facility's re-supply quantities? a) The facility itself (pull) b) The facility at the higher level (push/topping up) c) Other (explain in the comments section)	

Please circle answer(s)	Comments
21. How are the facility's re-supply quantities determined? a) Formula (describe in comment space) b) Higher level facility determines c) Other means (describe in comments section)	
22. Which data elements are used to calculate the facility's re-supply quantities? (Circle all that apply) a) Beginning of reporting period stock level b) End of reporting period stock level c) Quantity received d) Quantity dispensed e) Losses and adjustments f) Other (specify in the comments section)	
23. How did you learn how to complete the forms used at this facility? a) During a logistics training b) On the job training c) On the job d) Other (specify in the comments section)	
24. Who is responsible for transporting commodities to your facility? a) This facility collects b) The higher level facility delivers c) Other (explain in the comments section)	
25. What mode of transportation is most often used? a) Public transportation b) Facility-managed vehicle c) Private, hired vehicle d) On foot e) Other (specify in comments section)	
26. When did you conduct your last supervisory visit? a) Within the last month b) Within the last 3 months c) Within the last 6 months d) Other (explain in comments section) e) Never f) Not Applicable	

Please circle answer(s)	**Comments**
27. When did you receive your last supervisory visit? a) Within the last month b) Within the last 3 months c) Within the last 6 months d) Other (explain in comments section) e) Never *(go to question 30)* f) Not Applicable *(go to question 30)*	
28. Who conducted the supervisory visit that you received? (specify position of the person) _____	
29. What was done during the supervisory visit you received? (Circle all that apply) a) Supplies checked b) Stock cards checked c) Expired stock removed d) LMIS reports checked e) OJT/coaching f) Other (explain in the comments section)	

30. Please list all the HIV test kits that are managed by this facility:

Type	Brand

31. STOCK STATUS TABLE

Note the established minimum months of stock and maximum months of stock for fully supplied products in this facility and the time between orders.

Complete the following table for authorized products only using the following guidelines:

1. Enter all of the authorized products that will be studied.
2. Enter the units of count for each product (e.g. cycles, vials, tablets, pieces, etc).
3. Identify which products are managed by this facility by answering yes or no for each product.
4. Enter the total consumption or issues for the last 6 months. If less than 6 months of data available, enter data from as many months as possible.
5. Record the number of months the total consumption was based on (should be 6 in most cases)
6. Calculate the average monthly consumption in units of count for each product (column 4/column 5).
7. Record usable stock on hand based on a physical inventory of each product.
8. Record usable stock on hand based on stock ledgers or stock cards for each product. If there are no stock cards write "no stock cards".
9. Calculate months of stock on hand for each product. To calculate this, divide column 7 by column 6 (usable stock on hand from physical inventory/average monthly consumption).
10. If the number in column is less than the minimum number of months of stock recorded above the table, indicate whether or not an order has been placed with a yes or no. If the stock on hand is above the minimum, write "NA".
11. Enter the total amount of expired quantities of products that are on the shelf or anywhere inside the storeroom for each product.
12. Indicate the stock status as (–) if the months of stock on hand falls below the minimum and no order has been placed, (+) if months of stock on hand fall above the maximum and (=) if the months of stock falls between the minimum and the maximum.

Minimum: _____months   Maximum: _____months   Order interval: _____months

Product	Units of Count	Product managed by this facility? Y/N	Total consumption or issues last 6 months	Number of months	Average Monthly Consumption (col 4/col 5)	Usable stock on hand		Months of Stock on hand (col 7/col 6)	If Column 9 is less than min, has order been placed Y/N	Expired products	Stock Status (+,-,=) or comments
						From physical inventory	From stock ledger or stock cards				
1	2	3	4	5	6	7	8	9	10	11	12

32. STOCKOUT ASSESSMENT TABLE
Review the stock cards for the last 6 months to identify if any products stocked out. Alternatively, ask knowledgeable staff to identify if any products have stocked out over the past 6 months.

For all products that are both checked as available products and had a stockout in the last 6 months complete the following table:
¨**Note**: *It may be necessary to use more than one line per product in the table as, for example, there may have been 3 different stockouts of Depo-Provera in the last six months.*

1. Enter any products managed by this facility from the Stock Status Table that had a stock out in the last 6 months.
2. Record if there is a stock card available that has been updated within the last 6 months by answering yes or no.
3. Enter whether there was a stock out at the time of the visit for each product.
4. Enter the date (or estimated date) the stock out began.
5. Enter the date the stock out ended. If the stockout is on-going on the day of the visit, enter "on-going."
6. Calculate the duration of the stock out (add the number of days between column 4 and column 5). If the product is stocked out on the day of the visit, calculate the duration up to that day.
7. Check *column 7* if the date of the stockout has been taken from the stock cards or other logistics records.
8. Check *column 8* if an informant has estimated the date of the stockout.
9. Enter the reason for the stockout. Please use the following codes:

Reason for stockout:
1=Could not go to pick up the products
2=Higher level facility did not send enough products
3=Higher level facility did not send products in time
4=Increase in consumption
5=Did not request the right amount
6=Did not request products at the right time
7=Other reasons and state the reason

Authorized Products	Stock card available and updated (Y/N)	Stockout at time of visit (Y/N)	Stockout start date	Stockout end date	Duration of stock out (days between col. 5- col. 4)	Source of information		Reason for stockout (see list above)
						stock cards or other records	Informants knowledge	
1	2	3	4	5	6	7	8	9

33. STOCK DATA QUALITY TABLE:

If no stock cards are available, these tables can not be completed and you should skip to question 34.

A. Usable stock on hand on the day of the visit
1. Select a list of 10 products as a sample for calculating these indicators. This list can be determined before the visit and for the country as a whole, from the input of study team members prior to the fieldwork.
2. For column 2 below, copy from stock status table (table 31-column 7).
3. For column 3 below, copy from stock status table (table 31-column 8).
4. Note the quantity of products received that are not accounted for on the stock cards.
5. Note the quantity of products issued that are not accounted for on the stock cards.
6. Adjust the stock record (col. 3) by adding the unaccounted receipts (col. 4) and by subtracting the unaccounted issues (col. 5).
7. Calculate the percentage of discrepancy by subtracting the physical inventory count (column 2) from the adjusted stock record (column 6), divided by column 2, and multiplying by 100.

Note the reasons for discrepancy.

Method/Brand/ Product	Usable Stock on Hand (Day of Visit)						Reasons for Discrepancy
	From Physical Inventory	From Stock Ledger or Stock Cards	Count of unposted receipts	Count of unposted issues	Adjusted stock record (col3+col4-col 5)	% Discrepancy (col.6-col.2/col.2) *100	
1	2	3	4	5	6	7	8

B) Usable stock on hand at time of most recent LMIS report

1. List the same products as in table 33A in column 1.
2. Get the most recent LMIS report showing the selected products and write in column 2 the stock on hand from the LMIS report.
3. Write in column 3 the quantity of usable stock on hand from the stock records from the time of the LMIS report chosen.
4. Calculate the discrepancy by subtracting quantities of stock on hand from the LMIS report from quantities of stock on hand from stock records (from time of LMIS report)/ quantities of stock on hand from stock records times 100.

Note the reasons for discrepancy.

Method/Brand/ Product	Usable Stock on Hand (at time of most recent LMIS report)			
	According to Most Recent LMIS Report	From Stock Ledger or Stock Cards from time of LMIS Report	% Discrepancy ((col.3-col.2)/col.2) *100	Reasons for Discrepancy
1	2	3	4	5

34. FORECAST ACCURACY AND ORDER FILL RATE TABLE

1. Enter all of the authorized products that are included in this assessment and that are managed by this facility. This should include all products listed in table 31 and marked "yes" in column 3 of table 31.
2. Enter the amount forecasted for the last order period.
3. Enter the quantity consumed during last order period.
4. Calculate the forecast accuracy by subtracting the quantity forecasted for the last order period from the quantity consumed during the last order period, divided by the quantity forecasted, multiplied times 100.
5. Enter the quantity ordered for the last order period.
6. Enter the date order was placed.
7. Enter the quantity received in last order.
8. Enter the date order was received.
9. Calculate order fill rate by subtracting the quantity received for the last order period from quantity ordered for the last order period, divided by the quantity ordered for last order period, multiplied times 100.
10. Calculate the order lead time as the number of days between the date the order was placed and when it was received.

Write comments on the number of orders that arrived on schedule for the last 4 orders.

Method/Brand/ Product	Quantity forecasted for last order period	Quantity consumed during last order period	Forecast accuracy (col 2-col3)/col3 *100	Quantity ordered for last order period	Date order placed	Quantity received in last order/ procurement	Date order received	Percent diff. betw. quant. ordered & quant. received ((col.7-col.5)/ col.5) *100	Order lead time (days between col. 6 and col.8)	Comments: How many of your last 4 orders/procure ments were received according to schedule?
1	2	3	4	5	6	7	8	9	10	11

35. STORAGE/WAREHOUSE CONDITIONS TABLE

Items 1-13 should be assessed for all facilities. A table should be filled out for each storage area housing products being assessed. Please specify the types of products being assessed in the storage area by circling the category (ies) of products below.

Place a check mark in the appropriate column based on visual inspection of the storage facility, noting any relevant observations in the Comments column*. To qualify as "yes," all products and cartons must meet the criteria for each item*.

     Essential Drugs   Contraceptives   Vaccines   HIV test kits   STI drugs   TB/Leprosy

No	Description	Yes	No	N/A	Comments
1.	Products that are ready for distribution are arranged so that identification labels and expiry dates and/or manufacturing dates are visible.				
2.	Products are stored and organized in a manner accessible for First-Expiry / First-Out (FEFO) counting and general management.				
3.	Cartons and products are in good condition, not crushed due to mishandling. If opened cartons, products are not wet or cracked due to heat/radiation (fluorescent lights in the case of condoms)				
4.	The facility makes it a practice to separate damaged and/or expired products from good products and remove them from inventory.				
5.	Products are protected from direct sunlight at all times of the day and during all seasons.				
6.	Cartons and products are protected from water and humidity during all seasons.				
7.	Storage area is visually free from harmful insects and rodents. (Check the storage area for traces of rodents (droppings) or insects).				
8.	Storage area is secured with a lock and key, but accessible during normal working hours, with access limited to authorized personnel.				
9.	Products are stored at the appropriate temperature during all seasons according to product temperature specifications.				
10.	All hazardous waste (e.g., needles, toxic materials) is properly disposed of and non-accessible to non-medical personnel.				
11.	Roof is maintained in good condition to avoid sunlight and water penetration at all times.				
12.	Storeroom is maintained in good condition (e.g. clean, all trash removed, shelves are sturdy, boxes are organized).				

No.	Description	Yes	No	N/A	Comments
13.	The current space and organization is sufficient for existing products and reasonable expansion (i.e., receipt of expected product deliveries for the foreseeable future).				

The additional standards below can be applied to any facility large enough to require stacking of multiple boxes.

No.	Description	Yes	No	N/A	Comments
14.	Products are stacked at least 10 cm (4 inches) off the floor.				
15.	Products are stacked at least 30 cm (1 foot) away from the walls and other stacks.				
16.	Products are stacked no more than 2.5 meters (8 feet) high.				
17.	Fire safety equipment is available and accessible (any item identified as being used to promote fire safety should be considered).				
18.	Products are stored separately from insecticides and chemicals.				

Additional guidelines for specific questions:
Item 2: In noting proper product arrangement, the shelf life of the different products should be considered.
Item 3: Cartons should be checked to determine whether they are smashed due to mishandling. The conditions of the products inside opened or damaged cartons should also be examined to see if they are wet, cracked open due to heat/radiation (e.g. because of fluorescent lights in the case of condoms) or crushed.
Item 4: The discarding of damaged or expired products should be conducted according to the facility's procedures (which may differ from one facility to another). Please specify if procedures exist and note what they are.
Item 7: It is important to check the storage area for traces of rodents (droppings) or insects harmful to the products.
Item 8: This refers to either a warehouse secured with a lock or to a cabinet with a key in a clinic.
Item 17: Fire safety equipment does not have to meet international standards. Any item identified as being used to promote fire safety (e.g. water bucket, sand) should be considered.

No.	Question (and instructions)	Response (and skip instructions)
If the study team is not studying a cold chain logistics system, skip to question 41		
36	Do you have a functioning refrigerator(s) to store vaccines and/or HIV test kits?	1. yes number___ 2. no **(go to Q40)** 3. Not applicable **(go to Q40)**
37	Write down the actual temperature by looking at the internal thermometer inside the fridge (ideally temperature should be between 0 and +8 degrees centigrade) *(note if thermometer broken or missing)*	Temperature (Centigrade) _____
38	Are refrigerators located away from any surrounding objects?	1. yes 2. no
39	Is the temperature chart up-to-date? (in order to be up-to-date, there has to be an entry for the day of the visit)	1. yes 2. no
40	Do you have a supply of paraffin or LPG for cold chain and sterilization purposes?	1. yes 2. no

41	Are there any commodities you always run out of before re-supply? List the three most frequent.	1. _____ 2. _____ _____ 3. _____
42	Are there any commodities you always have a surplus of before re-supply? List the three most frequent.	1. _____ 2. _____ _____ 3. _____
43	What could be done to ensure the regular supply of products?	1. _____ 2. _____ 3. _____
44	Aside from "more staff" and "salary issues", what kind of support could be provided to help you do your job more effectively?	1. _____ 2. _____ 3. _____
Ask the person/people you interviewed if they have any questions for you.		

Thank the person/people who talked with you. Reiterate how they have helped the program achieve it's objectives and assure them that the results will be used to develop improvements in logistics system performance

45. End time of interview: _____   46. Total time of interview: _____

LOGISTICS INDICATOR ASSESSMENT TOOL (LIAT)[1]

INTERVIEWER'S GUIDE

Questions #1-4	Note the date, your name and the names of other interviewers, the time you begin the interview and the location of the facility.
Questions #5-8	Circle the type of facility where the survey is conducted and write in the name of the facility.
Question #9	Record the total number of staff at the facility. For warehouses, include staff of the warehouse only, even if it is linked to an SDP such as a district hospital. For SDPs include all staff in the facility, not only those working on logistics.
Question #10	List the titles of all the members of the staff interviewed at the site, how long they held the position and if they have received any training in logistics. Ask about the type of logistics training received by the staff in each product category (e.g., if speaking to the person in charge of immunization ask specifically if he/she has received formal training in cold chain and vaccine logistics).
Questions #11-29	Be sure to read each question exactly as written. If the respondent does not understand the question, you may explain it in your own words or translate to local dialect. If the answer given is not clear or does not clearly fit any response category, use the comment section to provide additional useful information.
Question #30	List all the HIV/test kits by type and brand. Please note if not applicable.
Tables #31-34	Should be filled for each product, following the instructions preceding each table in the questionnaire.
Table 35	Should be filled for each storage area housing products being assessed, following the instructions preceding the table.
Questions #36-40	Answer these questions only if the study team is looking at a cold chain logistics system.
Questions #41-44	Be sure to answer each question exactly as written. If the respondent does not understand the question, you may explain it in your own words or translate to local dialect. If the answer given is not clear or does not clearly fit any response category, use the comments section to provide additional useful information.
Questions #45-46	Record the ending time of the interview and calculate the total time of the interview using the beginning time recorded at the beginning of the questionnaire.

---

[1] DELIVER/JSI, 2002.

QUESTIONNAIRE
FOR SERVICE DELIVERY POINTS, DISTRICT/REGIONAL FACILITIES AND CENTRAL WAREHOUSES

Introduce yourself and all members of the team, including titles/positions. Present the objectives of this assessment and how this interview will help the team to achieve the objectives.

State your objectives here

Explain how the team will conduct the interview, invite relevant interviewees to join the group and begin.

Beginning time of Interview _____ | End time of Interview _____

1. Date _____ 2. Interviewer(s) _____

3. Region/Province _____ **4.** District _____

5. Type of facility: (Circle all that apply)
a) Urban/Rural
b) Public/Private
c) Warehouse/Service Delivery Point **(if warehouse go to question 6, if SDP go to question 7)**

6. If warehouse:
a) Central
b) Regional/Provincial
c) District/Zonal

7. If SDP:
a)   Hospital
b)   Clinic
c)   Dispensary
d)   Health post
e)   Health worker

8. Name of the facility: _____

9. Total number of staff at facility: _____

10. Respondents interviewed at this site:

    Title   Length in current position   **Received training in logistics**
                                                                   (**specify product categories & dates**)

a) _____years/months   _____

b) _____years/months   _____

c) _____years/months   _____

d) _____years/months   _____

e) _____years/months   _____

f) _____years/months   _____

g) _____years/months   _____

*** Logistics includes the following functions: Ordering, Receiving supplies, Inventory management, and supervision. If speaking to the person in charge of EPI, ask specifically "Have you received formal training for cold chain and vaccine logistics?"*

11. Do you use the following logistics forms to manage health products?	
a) Stock cards/records	*If no, go to question 16*
b) LMIS reports (including requisition and issue vouchers, consumption, transactions)	

Please circle answer(s)	**Comments**

12. How is the information on these forms used? (Circle all that apply)
  a) Calculating consumption
  b) Calculating needs
  c) Reporting use to the higher level
  d) Requesting supplies from the higher level
  e) Other, please explain in comments section	

13. If LMIS reports are used, are they sent to the higher level? a) Yes b) No *(go to question 15)* c) Don't know *(go to question16)* d) Not applicable *(go to question16)*	

Please circle answer(s)	Comments
14. How often are these forms sent to the higher level? a) Monthly b) Quarterly c) Semi-annually d) Annually e) Other (please specify in the comments section) f) Not Applicable	
15. How often are you supposed to send these forms to the higher level? a) Monthly b) Quarterly c) Semi-annually d) Annually e) Other (please specify in the comments section) f) Not Applicable	
16. How many facilities should send reports to this facility? _____ **(if zero, go to Q18)**	
17. Provide an approximate number of facilities that send these reports according to schedule. _____	
18. How many times have you placed an order or submitted a procurement request in the last year? a) None b) 0-3 times c) 3-6 times d) more than 6 times a year (specify reasons in the comments section)	
19. How often are you supposed to place orders or submit a procurement request? a) Monthly b) Quarterly c) Semi-annually d) Annually e) Other (specify in the comments section) f) Not Applicable	
20. Who determines this facility's re-supply quantities? a) The facility itself (pull) b) The facility at the higher level (push/topping up) c) Other (explain in the comments section)	

Please circle answer(s)	Comments
21. How are the facility's re-supply quantities determined? a) Formula (describe in comment space) b) Higher level facility determines c) Other means (describe in comments section)	
22. Which data elements are used to calculate the facility's re-supply quantities? (Circle all that apply) a) Beginning of reporting period stock level b) End of reporting period stock level c) Quantity received d) Quantity dispensed e) Losses and adjustments f) Other (specify in the comments section)	
23. How did you learn how to complete the forms used at this facility? a) During a logistics training b) On the job training c) On the job d) Other (specify in the comments section)	
24. Who is responsible for transporting commodities to your facility? a) This facility collects b) The higher level facility delivers c) Other (explain in the comments section)	
25. What mode of transportation is most often used? a) Public transportation b) Facility-managed vehicle c) Private, hired vehicle d) On foot e) Other (specify in comments section)	
26. When did you conduct your last supervisory visit? a) Within the last month b) Within the last 3 months c) Within the last 6 months d) Other (explain in comments section) e) Never f) Not Applicable	

Please circle answer(s)	Comments
27. When did you receive your last supervisory visit? a) Within the last month b) Within the last 3 months c) Within the last 6 months d) Other (explain in comments section) e) Never *(go to question 30)* f) Not Applicable *(go to question 30)*	
28. Who conducted the supervisory visit that you received? (specify position of the person) _____	
29. What was done during the supervisory visit you received? (Circle all that apply) a) Supplies checked b) Stock cards checked c) Expired stock removed d) LMIS reports checked e) OJT/coaching f) Other (explain in the comments section)	

30. Please list all the HIV test kits that are managed by this facility:

Type	Brand

31. STOCK STATUS TABLE

Note the established minimum months of stock and maximum months of stock for fully supplied products in this facility and the time between orders.

Complete the following table for authorized products only using the following guidelines:

1. Enter all of the authorized products that will be studied.
2. Enter the units of count for each product (e.g. cycles, vials, tablets, pieces, etc).
3. Identify which products are managed by this facility by answering yes or no for each product.
4. Enter the total consumption or issues for the last 6 months. If less than 6 months of data available, enter data from as many months as possible.
5. Record the number of months the total consumption was based on (should be 6 in most cases)
6. Calculate the average monthly consumption in units of count for each product (column 4/column 5).
7. Record usable stock on hand based on a physical inventory of each product.
8. Record usable stock on hand based on stock ledgers or stock cards for each product. If there are no stock cards write "no stock cards".
9. Calculate months of stock on hand for each product. To calculate this, divide column 7 by column 6 (usable stock on hand from physical inventory/average monthly consumption).
10. If the number in column is less than the minimum number of months of stock recorded above the table, indicate whether or not an order has been placed with a yes or no. If the stock on hand is above the minimum, write "NA".
11. Enter the total amount of expired quantities of products that are on the shelf or anywhere inside the storeroom for each product.
12. Indicate the stock status as (–) if the months of stock on hand falls below the minimum and no order has been placed, (+) if months of stock on hand fall above the maximum and (=) if the months of stock falls between the minimum and the maximum.

Minimum: _____months   Maximum: _____months   Order interval: _____months

Product	Units of Count	Product managed by this facility? Y/N	Total consumption or issues last 6 months	Number of months	Average Monthly Consumption (col 4/col 5)	Usable stock on hand		Months of Stock on hand (col 7/col 6)	If Column 9 is less than min, has order been placed Y/N	Expired products	Stock Status (+,-,=) or comments
						From physical inventory	From stock ledger or stock cards				
1	2	3	4	5	6	7	8	9	10	11	12

D-476

32. STOCKOUT ASSESSMENT TABLE
Review the stock cards for the last 6 months to identify if any products stocked out. Alternatively, ask knowledgeable staff to identify if any products have stocked out over the past 6 months.

For all products that are both checked as available products and had a stockout in the last 6 months complete the following table:
¨**Note***: It may be necessary to use more than one line per product in the table as, for example, there may have been 3 different stockouts of Depo-Provera in the last six months.*

1. Enter any products managed by this facility from the Stock Status Table that had a stock out in the last 6 months.
2. Record if there is a stock card available that has been updated within the last 6 months by answering yes or no.
3. Enter whether there was a stock out at the time of the visit for each product.
4. Enter the date (or estimated date) the stock out began.
5. Enter the date the stock out ended. If the stockout is on-going on the day of the visit, enter "on-going."
6. Calculate the duration of the stock out (add the number of days between column 4 and column 5). If the product is stocked out on the day of the visit, calculate the duration up to that day.
7. Check *column 7* if the date of the stockout has been taken from the stock cards or other logistics records.
8. Check *column 8* if an informant has estimated the date of the stockout.
9. Enter the reason for the stockout. Please use the following codes:

Reason for stockout:
1=Could not go to pick up the products
2=Higher level facility did not send enough products
3=Higher level facility did not send products in time
4=Increase in consumption
5=Did not request the right amount
6=Did not request products at the right time
7=Other reasons and state the reason

Authorized Products	Stock card available and updated (Y/N)	Stockout at time of visit (Y/N)	Stockout start date	Stockout end date	Duration of stock out (days between col. 5- col. 4)	Source of information		Reason for stockout (see list above)
						stock cards or other records	Informants knowledge	
1	2	3	4	5	6	7	8	9

33. STOCK DATA QUALITY TABLE:

If no stock cards are available, these tables can not be completed and you should skip to question 34.

A. Usable stock on hand on the day of the visit
1. Select a list of 10 products as a sample for calculating these indicators. This list can be determined before the visit and for the country as a whole, from the input of study team members prior to the fieldwork.
2. For column 2 below, copy from stock status table (table 31-column 7).
3. For column 3 below, copy from stock status table (table 31-column 8).
4. Note the quantity of products received that are not accounted for on the stock cards.
5. Note the quantity of products issued that are not accounted for on the stock cards.
6. Adjust the stock record (col. 3) by adding the unaccounted receipts (col. 4) and by subtracting the unaccounted issues (col. 5).
7. Calculate the percentage of discrepancy by subtracting the physical inventory count (column 2) from the adjusted stock record (column 6), divided by column 2, and multiplying by 100.
Note the reasons for discrepancy.

Method/Brand/ Product	Usable Stock on Hand (Day of Visit)						Reasons for Discrepancy
	From Physical Inventory	From Stock Ledger or Stock Cards	Count of unposted receipts	Count of unposted issues	Adjusted stock record (col3+col4-col 5)	% Discrepancy (col.6-col.2/col.2) *100	
1	2	3	4	5	6	7	8

B) Usable stock on hand at time of most recent LMIS report

1. List the same products as in table 33A in column 1.
2. Get the most recent LMIS report showing the selected products and write in column 2 the stock on hand from the LMIS report.
3. Write in column 3 the quantity of usable stock on hand from the stock records from the time of the LMIS report chosen.
4. Calculate the discrepancy by subtracting quantities of stock on hand from the LMIS report from quantities of stock on hand from stock records (from time of LMIS report)/ quantities of stock on hand from stock records times 100.

Note the reasons for discrepancy.

Method/Brand/ Product	Usable Stock on Hand (at time of most recent LMIS report)			
	According to Most Recent LMIS Report	From Stock Ledger or Stock Cards from time of LMIS Report	% Discrepancy ((col.3-col.2)/col.2) *100	Reasons for Discrepancy
1	2	3	4	5

34. FORECAST ACCURACY AND ORDER FILL RATE TABLE

1. Enter all of the authorized products that are included in this assessment and that are managed by this facility. This should include all products listed in table 31 and marked "yes" in column 3 of table 31.
2. Enter the amount forecasted for the last order period.
3. Enter the quantity consumed during last order period.
4. Calculate the forecast accuracy by subtracting the quantity forecasted for the last order period from the quantity consumed during the last order period, divided by the quantity forecasted, multiplied times 100.
5. Enter the quantity ordered for the last order period.
6. Enter the date order was placed.
7. Enter the quantity received in last order.
8. Enter the date order was received.
9. Calculate order fill rate by subtracting the quantity received for the last order period from quantity ordered for the last order period, divided by the quantity ordered for last order period, multiplied times 100.
10. Calculate the order lead time as the number of days between the date the order was placed and when it was received.

Write comments on the number of orders that arrived on schedule for the last 4 orders.

Method/Brand/ Product	Quantity forecasted for last order period	Quantity consumed during last order period	Forecast accuracy (col 2- col3)/col3 *100	Quantity ordered for last order period	Date order placed	Quantity received in last order/ procurement	Date order received	Percent diff. betw. quant. ordered & quant. received ((col.7- col.5)/ col.5) *100	Order lead time (days between col. 6 and col.8)	Comments: How many of your last 4 orders/procure ments were received according to schedule?
1	2	3	4	5	6	7	8	9	10	11

35. STORAGE/WAREHOUSE CONDITIONS TABLE

Items 1-13 should be assessed for all facilities. A table should be filled out for each storage area housing products being assessed. Please specify the types of products being assessed in the storage area by circling the category (ies) of products below.

Place a check mark in the appropriate column based on visual inspection of the storage facility, noting any relevant observations in the Comments column. *To qualify as "yes," all products and cartons must meet the criteria for each item*.

Essential Drugs   Contraceptives   Vaccines   HIV test kits   STI drugs   TB/Leprosy

No	Description	Yes	No	N/A	Comments
1.	Products that are ready for distribution are arranged so that identification labels and expiry dates and/or manufacturing dates are visible.				
2.	Products are stored and organized in a manner accessible for First-Expiry / First-Out (FEFO) counting and general management.				
3.	Cartons and products are in good condition, not crushed due to mishandling. If opened cartons, products are not wet or cracked due to heat/radiation (fluorescent lights in the case of condoms)				
4.	The facility makes it a practice to separate damaged and/or expired products from good products and remove them from inventory.				
5.	Products are protected from direct sunlight at all times of the day and during all seasons.				
6.	Cartons and products are protected from water and humidity during all seasons.				
7.	Storage area is visually free from harmful insects and rodents. (Check the storage area for traces of rodents (droppings) or insects).				
8.	Storage area is secured with a lock and key, but accessible during normal working hours, with access limited to authorized personnel.				
9.	Products are stored at the appropriate temperature during all seasons according to product temperature specifications.				
10.	All hazardous waste (e.g., needles, toxic materials) is properly disposed of and non-accessible to non-medical personnel.				
11.	Roof is maintained in good condition to avoid sunlight and water penetration at all times.				
12.	Storeroom is maintained in good condition (e.g. clean, all trash removed, shelves are sturdy, boxes are organized).				

No.		Description	Yes	No	N/A	Comments
13.		The current space and organization is sufficient for existing products and reasonable expansion (i.e., receipt of expected product deliveries for the foreseeable future).				

The additional standards below can be applied to any facility large enough to require stacking of multiple boxes.

No.	Description	Yes	No	N/A	Comments
14.	Products are stacked at least 10 cm (4 inches) off the floor.				
15.	Products are stacked at least 30 cm (1 foot) away from the walls and other stacks.				
16.	Products are stacked no more than 2.5 meters (8 feet) high.				
17.	Fire safety equipment is available and accessible (any item identified as being used to promote fire safety should be considered).				
18.	Products are stored separately from insecticides and chemicals.				

Additional guidelines for specific questions:
Item 2: In noting proper product arrangement, the shelf life of the different products should be considered.
Item 3: Cartons should be checked to determine whether they are smashed due to mishandling. The conditions of the products inside opened or damaged cartons should also be examined to see if they are wet, cracked open due to heat/radiation (e.g. because of fluorescent lights in the case of condoms) or crushed.
Item 4: The discarding of damaged or expired products should be conducted according to the facility's procedures (which may differ from one facility to another). Please specify if procedures exist and note what they are.
Item 7: It is important to check the storage area for traces of rodents (droppings) or insects harmful to the products.
Item 8: This refers to either a warehouse secured with a lock or to a cabinet with a key in a clinic.
Item 17: Fire safety equipment does not have to meet international standards. Any item identified as being used to promote fire safety (e.g. water bucket, sand) should be considered.

No.	Question (and instructions)	Response (and skip instructions)
If the study team is not studying a cold chain logistics system, skip to question 41		
36	Do you have a functioning refrigerator(s) to store vaccines and/or HIV test kits?	1. yes number___ 2. no **(go to Q40)** 3. Not applicable **(go to Q40)**
37	Write down the actual temperature by looking at the internal thermometer inside the fridge (ideally temperature should be between 0 and +8 degrees centigrade) *(note if thermometer broken or missing)*	Temperature (Centigrade) _____
38	Are refrigerators located away from any surrounding objects?	1. yes 2. no
39	Is the temperature chart up-to-date? (in order to be up-to-date, there has to be an entry for the day of the visit)	1. yes 2. no
40	Do you have a supply of paraffin or LPG for cold chain and sterilization purposes?	1. yes 2. no

41	Are there any commodities you always run out of before re-supply? List the three most frequent.	1. _____ 2. _____ _____ 3. _____
42	Are there any commodities you always have a surplus of before re-supply? List the three most frequent.	1. _____ 2. _____ _____ 3. _____
43	What could be done to ensure the regular supply of products?	1. _____ 2. _____ 3. _____
44	Aside from "more staff" and "salary issues", what kind of support could be provided to help you do your job more effectively?	1. _____ 2. _____ 3. _____
Ask the person/people you interviewed if they have any questions for you.		

Thank the person/people who talked with you. Reiterate how they have helped the program achieve it's objectives and assure them that the results will be used to develop improvements in logistics system performance

45. End time of interview: _____   46. Total time of interview: _____

LOGISTICS SYSTEM ASSESSMENT TOOL (LSAT)[1]

SECTION I-ORGANIZATION
Attach a copy of the organizational chart describing the logistics personnel for the supply chain being assessed.

Organizational Context				
	Yes	**No**	**NA**	**Comments**
1.1 Is there a logistics management unit at the national level?				
If no, please note in the comments section(questions 1.2 a-i) which departments or positions are responsible for each logistics task				
1.2 Is the logistics management unit fully responsible for the following activities:				
a) Logistics Management Information System?				
b) Forecasting quantities needed?				
c) Procurement?				
d) Inventory management, storage and distribution?				
e) Product selection?				
f) Determining the organizational structure and processes?				
g) Staffing of logistics positions?				
h) Budgeting for the logistics system?				
i) Supervision and logistic staff development?				
1.3 Are there documented guidelines for:				
a) Logistics information management systems?				
b) Forecasting quantities needed?				
c) Procurement?				
d) Inventory management, storage and distribution?				
e) Product selection?				
f) Staffing of logistics positions?				
g) Budgeting for the logistics system?				
h) Supervision and staff development?				

---

[1] DELIVER/JSI, 2002.

1.4	Is there a dedicated logistics officer-in-charge?				

If no, skip to question # 1.6

1.5	Does the logistics officer(s)-in-charge have the same level of authority as other functional unit heads for decision-making?				

1.6	What are the mechanisms used to coordinate key logistics tasks among those responsible for logistics?

☐ None   → *skip to question #1.8*   ☐ Formal meetings
☐ Joint work plans   ☐ Written communications
☐ Department meetings   ☐ Other_____

1.7	Are these coordination mechanisms effective?

1.8	How many of the positions with key logistics tasks are currently filled? If not filled, why?

1.9	Describe the relationships among key stakeholders, including donors, other cooperating agencies, government units, other relevant entities and other supply chains.

		Yes	No	NA	Comments
1.10	Is there a logistics system strategy or improvement plan that guides the activities of the system?				

If no, skip to question #1.12

1.11	If yes, describe the strategy/plan.

1.12	What issues outside the supply chain impact the functioning of the supply chain?

(Note: Include major political, cultural, or economic factors such as political events, labor disputes, etc.)

SECTION II-LOGISTICS MANAGEMENT INFORMATION SYSTEM (LMIS)

Basic elements of an LMIS					
		Yes	No	NA	Comments
2.1	Is there a logistics management information system?				

If yes, go to question 2.3

2.2	If no, is logistics information collected through another information system (e.g. HMIS)? Describe briefly.

		Yes	No	NA	Comments
2.3	Does the information system include:				
a)	Stock keeping records (e.g., inventory control cards, bin cards, stock registers) at all levels?				
b)	Requisition and issue records (e.g., bills of lading, shipping records, requisition/issue vouchers) at all levels?				

		Yes	No	NA	Comments
c)	Dispensed-to-user records at service delivery points?				
d)	Summaries of consumption data at levels above service delivery points (e.g., districts, regions, central, etc)?				
2.4	Do information system reports at all levels of the system show:				
a)	The inventory balance (stock on hand)?				
b)	Quantity dispensed or issued during a specified reporting period?				
c)	Losses and adjustments?				

If no, skip to question #2.6

2.5 a) b)	What indicators does the information system track? (e.g. stockout rate, percentage of reporting, rational prescribing practices, etc) Who tracks these indicators?

2.6	How often are logistics reports supposed to be sent to the next highest level of the system? ☐ Monthly ☐ Quarterly ☐ Semi-annually ☐ Annually ☐ Other_____

2.7	Do facilities follow this reporting schedule? Map the report flow.

2.8	How do managers monitor reporting rates and follow up to obtain missing logistics reports?

2.9 a) b) c)	Are information system records reconciled against physical inventories at each level? How is this done? How often? ☐ Monthly, ☐ Quarterly, ☐ Semi-annually, ☐ Annually ☐Other_____

2.10	What is the approximate percentage of information system reports received in time to be used for logistics decisions at each level of the system? Central _____, Regional _____, District _____

		Yes	No	NA	Comments
2.11	Is the information system automated at:				
a)	the central level				
b)	the regional level				
c)	the district level				
d)	the service delivery point				

If no to questions # 2.11 a - d, skip to 2.13

2.12	Briefly describe the functions and processes that are automated.
2.13	How is logistics data recorded, managed and analyzed at each level?
2.14	Is external assistance provided to manage the information system?

Use of LMIS Information				
	Yes	**No**	**NA**	**Comments**
2.15 Is the information system used for monitoring and evaluating the program's performance?				
2.16 What decisions are based on information system reports? ☐ Forecasting ☐ Procurement ☐ Transport/Delivery ☐ Scheduling supervisory visits ☐ Other_____				
	Yes	**No**	**NA**	**Comments**
2.17 Are logistics data used at each level of the system as appropriate for:				
a) Continuous monitoring of stock balances?				
b) Calculating quantities for re-supply?				
2.18 a) What feedback mechanisms are in place to channel logistics information back to lower levels? b) How is information fed back? ☐ None ☐ Telephone call ☐ Reports ☐ Meetings ☐ Supervisory visit ☐ Other _____				
2.19 a) Are issues data or dispensed-to-user data cross-checked against other data sources? ☐ Service statistics ☐ Demographic statistics ☐ Survey data ☐ Field audit data ☐ Other_____ b) How often are they checked? ☐ Quarterly ☐ Semi-annually ☐ Annually ☐ Other_____ c) Who is responsible for this cross-checking?				
2.20 a) Is logistics information provided to appropriate decision-makers for logistics planning (e.g., Ministry of Health, Ministry of Finance, UNFPA, USAID, World Bank, NGOs)? ☐ Yes ☐ No b) What information is provided? c) By whom? d) How often? ☐ Monthly ☐ Quarterly ☐ Semi-annually ☐ Annually ☐ Other_____ e) How is the information used?				

SECTION III- PRODUCT SELECTION

National Drug Policy				
	Yes	No	NA	Comments
3.1 Is there a National Drug Policy document?				
If no, skip to question #3.4.				
3.2 a) When was it published? Attach a copy. Date_____ b) Who develops it? And how often is it updated? Who receives it? c) How is it used?				
	Yes	No	NA	Comments
3.3 Does the National Drug Policy contain written guidelines for donation of products?				
3.4 Is duty tax imposed on imported drugs or products?				
3.5 Are donated commodities exempt from duty tax?				
3.6 How are new drugs or products registered?				
3.7 Does the program have a written policy for maintaining continuity of brands and avoiding unnecessary duplication of interchangeable products (e.g. hormonal formulations of contraceptives and socially marketed products)?				
3.8 Is there an essential services package? If yes, what services are included?				

National Essential Drug List				
	Yes	No	NA	Comments
3.10 Is there a national essential drugs list?				
If no, skip to section IV.				
3.11 What categories of products does the list include? ☐ Family Planning ☐ STI ☐ HIV/AIDS ☐ Essential Drugs ☐ TB ☐ Malaria ☐ Vaccines ☐ Vitamin Supplements ☐ Other				
3.12 List all contraceptives that are available in the country and specify which ones are on the essential drug list.				
3.13 How many products, including contraceptives, does the list contain? (Provide a copy of the list)				
3.14 What are the criteria for inclusion of a product to the list?				

3.15					
a)	Is the national essential drugs list officially distributed to all levels of the system? □ Central □ Regional □ District □ Service delivery point				
b)	Is it used for product selection and ordering commodities? If yes, explain how it is used.				

SECTION IV-FORECASTING

National Level Forecast Preparation				
	Yes	**No**	**NA**	**Comments**
4.1 Are forecasts developed using:				
a) Logistics based data?				
b) Demographic data or disease prevalence/morbidity?				
c) Service statistics?				
4.2 If forecasts are prepared and updated using the most recent logistics data, do they include:				
a) Stock on hand?				
b) Dispensed-to-user data?				
c) Losses and adjustments?				
4.3 Are forecasts validated by comparing previous estimated consumption with actual consumption?				
4.4 What other factors are considered in the preparation of forecasts? (e.g., consolidating decentralized forecasts or quantifications, seasonal and regional variations, standard treatment guidelines, national essential drug list, etc.)				
4.5 Do forecasts take into account programmatic plans (e.g., expansion of service outlets, training, IEC or behavior change campaigns, other organizations, etc.)?				
4.6 Describe the forecasting process? a) Who initiates it? b) When it takes place? c) How long does the process take?				
4.7 a) Is technical assistance provided to develop correct forecasts? □ Yes □ No b) If yes, by whom?				
4.8 What is the role of regional or lower levels in the forecasting process?				
	Yes	**No**	**NA**	**Comments**
4.9 Are forecasts updated at least annually?				
4.10 Are forecasts prepared on a schedule coinciding with local budgeting and procurement cycles:				
a) Short-term (e.g., annual)?				
b) Long-term (e.g., three or more years)?				
4.11 Are forecasts costed out and incorporated into budget planning by the MOH and/or donors? Explain.				

4.12
a)   Are there funding shortfalls?   ☐ Yes   ☐ No
b)   If so, how are they resolved?

SECTION V-OBTAINING SUPPLIES/PROCUREMENT

Procurement planning				
5.1 Who is responsible for procurement planning/ordering and scheduling shipments (e.g., logistics unit, procurement unit) at appropriate levels?				
5.2 Describe the coordination between persons or unit(s) responsible for logistics activities and the procurement staff.				
	Yes	**No**	**NA**	**Comments**
5.3 Are short-term procurement plans based on forecasted needs?				
5.4 Do these procurement plans take into account the following logistics systems elements:				
a) Current inventory levels (stock on hand)?				
b) Losses and adjustments?				
c) Required order lead times of suppliers/donors?				
d) Established stock levels, if relevant?				
e) Shipment and handling schedules?				
f) The need for a safety stock?				
5.5 Are procurements limited to:				
a) Pre-qualified suppliers?				
b) Products on the national essential drugs list?				

5.6
a)   In general, are the correct amounts of all products procured and obtained in an appropriate time frame at all the following levels?   ☐ Central   ☐ Regional   ☐ District   ☐ Service delivery point
b)   Specify any products which do not arrive in a timely manner, in appropriate amounts and why.

5.7
a)   What are the procedures and time frames for ordering products from suppliers and donors?
b)   How do these take account of trade, regulatory, and currency restrictions?

5.8   What is done to monitor/manage the coordination of procurement plans among suppliers/donors?

5.9
a)   Is pipeline status regularly monitored so that procurement decisions can be made and actions can be initiated in time to avoid stockouts?
b)   Who does this and how?

	Yes	No	NA	Comments
5.10 Does the procurement unit or persons responsible for procurement:				
a) Write and issue tenders?				
b) Evaluate bids?				
c) Monitor supplier performance?				
5.11 Does the program have written procedures for ensuring that products received meet defined standards of quality?				
5.12 a) What are the procedures for quality assurance, who is responsible and how often are they done? b) Is there a procedure for recording and reporting complaints regarding product quality to suppliers?				
5.13 What other actions are carried out to ensure product quality?				

SECTION VI-INVENTORY CONTROL PROCEDURES

Stock Management				
6.1 Specify what type of inventory control system is used (e.g., push, pull, etc.) and describe the system.				
	Yes	**No**	**NA**	**Comments**
6.2 Are there guidelines and established policies for maximum and minimum stock levels at which full supply products should be maintained *(please note current max and min levels in comments section)*:				
a) At the central level of the supply chain?				
b) At the regional level of the supply chain?				
c) At the district level of the supply chain?				
d) At the service delivery point level of the supply chain?				
6.3 a) Are the inventory control guidelines for fully supplied products respected at all levels so that stock levels generally fall between maximum and minimum? b) If not, why not?				
6.4 a) Are stock levels (min, max) for full supply products reviewed periodically? b) Do they take into account changes in transport and information availability?				
6.5 How are products that cannot be maintained in full supply allocated? Specify by level. Central to regional? Regional to district? District to SDP?				

		Yes	No	NA	Comments
6.6	Are there written provisions for the redistribution of over-stocked supplies?				

6.7	How are stock imbalances handled by supervisors/managers at:
	Central? Regional? District? SDP?

		Yes	No	NA	Comments
6.8	Does the program have a policy of storing and issuing stock according to first expiry/first out inventory control procedures at all levels?				

If no, skip to question #6.11

6.9	In practice does the program manage and issue stock according to first expiry/first out inventory control procedures at all levels? Please describe.

6.10	Are damaged/expired products physically separated from inventory and removed from stock records at all levels? Note the approximate quantities of products expired within the last 2 years.

		Yes	No	NA	Comments
6.11	Does the program have a system for tracking product losses and				

	Yes	No	NA	Comments
6.17 Are there established procedures for placing emergency orders?				
6.18 How often are emergency orders filled at: Central? Regional? District? SDP?				

SECTION VII-WAREHOUSING AND STORAGE

Adequacy of Storage Capacity and Conditions				
	Yes	No	NA	Comments
7.1 Does the program have written guidelines for storage and handling of all products at all levels of the system? (E.g., manuals, posters, etc.)				
7.2 Are there written guidelines for disposal of sharps, bio-hazardous material, and other medical waste?				
7.3 Is there a policy that requires at least one physical inventory of all products per year at each storage facility?				
7.4 Are cold chain storage resources (e.g. fridge, paraffin/kerosene, temperature chart) available at all levels of the system, where appropriate?				
7.5 If applicable, how is the cold chain monitored to ensure that products are consistently maintained at appropriate temperatures?(check all that apply) ☐ Written guidelines ☐ Supervision ☐ Temperature log sheets ☐ Other				
7.6 Is the existing storage capacity adequate to handle the current quantities of products at: Central? Regional? District? SDP?				
7.7 Can the existing storage capacity handle all the quantities needed to ensure no stockouts at: The central level? The regional level? The district level? The SDP level?				
If yes to all, skip to question # 7.9				
7.8 If not, how does the program cope with inadequate storage space at: The central level? The regional level? The district level? The SDP level?				
	Yes	No	NA	Comments
7.9 Does the program have plans for meeting storage requirements for at least the next 5 years?				
7.10 Describe the program's plans for accommodating growth (e.g., infrastructure, distribution, etc.)?				
7.11 Specify storage conditions in need of improvement, if any (e.g., cleanliness, organization, temperature, building structure, etc.).				

Assuring Product Quality at the Storage Facilities				
	Yes	**No**	**NA**	**Comments**
7.12 Is there a procedure for recording complaints regarding product quality at all levels?				
If no, skip to question # 7.14				
7.13 If so, how are they addressed?				
7.14 Are visual quality assurance inspections of products conducted at the storage facility at the following levels				

Levels	Yes	No	How often?	Comments
Central?				
Regional?				
District?				
SDP?				

	Yes	**No**	**NA**	**Comments**
7.15 Are there written procedures or guidelines for destroying damaged and expired products?				
If no, skip to question # 7.17				
7.16 Describe the procedures/guidelines				
7.17 a) In practice, are damaged and expired products destroyed according to the program's disposal guidelines? b) If so, how often?				

SECTION VIII-TRANSPORT AND DISTRIBUTION

Distribution System				
	Yes	**No**	**NA**	**Comments**
8.1 In the program's budget is there a line item for: *(identify the program, e.g. MOH, FP program, Logistics Division, etc)*				
a) Vehicles?				
b) Fuel?				
c) Spare vehicle parts?				
d) Vehicle maintenance and repair?				
e) Per diems?				
f) Salaries for drivers?				
8.2 a) Are any of the above items supported by external funds? ☐ Yes ☐ No b) How much? By whom?				

8.3	If transport is externally funded, are there plans to phase-out/end this support?				
		Yes	**No**	**NA**	**Comments**
8.4	Do written procedures specify what type of distribution system is to be used to distribute products between each level?				

8.5	How are products delivered between each level of the system (include means of transportation)? Specify between which levels.				
		Yes	**No**	**NA**	**Comments**
8.6	Is there a documented distribution schedule for all levels?				

8.7	Which essential health products are distributed together (e.g., contraceptives, essential drugs, TB drugs, STI and HIV test kits and drugs, laboratory supplies, etc.)? Specify by level.
8.8 a) b)	Are there a sufficient number of functioning vehicles, with available petrol/drivers at appropriate levels to meet the desired distribution schedule? Are they regularly available for transport and other activities such as supervision?
8.9	Are vehicles used effectively for routine deliveries and emergency deliveries at all levels? Explain (e.g., maximum use of vehicle capacity, coordination of distribution routes, etc.)
8.10 a) b)	Are vehicles in running order? How is vehicle maintenance handled at the different appropriate levels?
8.11	Where are the vehicles (at what levels of the system)?
8.12	In general, are orders delivered as scheduled at the following levels: Central? Regional? District? SDP?
8.13	Is transportation outsourced at any level of the system? If yes, how effective has it been?

SECTION IX-ORGANIZATIONAL SUPPORT FOR LOGISTICS SYSTEM

Organizational Processes for Logistics				
	Yes	**No**	**NA**	**Comments**
9.1 Do personnel between these different levels communicate at least quarterly?				
a) Central level logistics staff and next level (e.g. region, province) staff				
b) Regional level of logistics staff with district level staff in their area				
c) District level logistics staff with the SDP level				
If no to question # 9.1 a-c, skip to question #9.3				

9.2	If yes to any of questions #9.1 a-c, describe the means of communication and what is usually covered.				
If no, skip to question # 9.6					
9.3	Describe these methods (e.g., regular meetings, phone calls, letters, radio, etc.). Which are most effective and why.				
9.4	Which of these communication methods used in *question # 9.3* are most effective and why?				
9.5	In the last year, have logistics functions been affected by waiting for decisions, approvals, information and/or guidance? If so, how?				
		Yes	**No**	**NA**	**Comments**
9.6	Is there an established mechanism for improving logistics practices or procedures (based on what is learned from supervisory visits, feedback, assessments, etc.) at the:				
a)	Central level?				
b)	Regional level?				
c)	District level?				
d)	SDP level?				
		Yes	**No**	**NA**	**Comments**
9.7	Are there written procedures and guidelines (e.g., manuals, job aids, standards) to help staff carry out their logistics responsibilities?				
If no, skip to question # 9.11					
9.8	List all procedures/guidelines that cover logistics responsibilities.				
		Yes	**No**	**NA**	**Comments**
9.9	Are the procedures and guidelines distributed to staff at the:				
a)	Central level?				
b)	Regional level?				
c)	District level?				
d)	SDP level?				
9.10	Do staff who manage commodities have a written job description that includes logistics responsibilities at the:				
a)	Central level?				
b)	Regional level?				
c)	District level?				
d)	SDP level?				
9.11	Do logistics staff have the tools and resources they need to do their jobs at all levels (e.g., job aids, forms, carbon paper, calculators, shelving, vehicles, funds for transport, etc.)? If not, which tools or resources are missing? Central? Regional? District? SDP?				
		Yes	**No**	**NA**	**Comments**
9.12	Is external assistance used to carry out management and supervision activities?				
If no, skip to question # 9.14					
9.13	If yes, describe the extent of the external assistance.				

Supervision (Individual Performance Management)				
9.14	Describe supervisory relationships by job position/title and by level. Indicate if any position receives supervision from more than one person or unit. Provide a chart if possible.			
		Yes	**No**	**NA**	**Comments**
9.15	Are supervisory responsibilities described in written job descriptions?				
9.16	Are there guidelines for how the supervisor conducts the supervisory visit (e.g., introductions, positive style of interaction, follow-up)?				
9.17	Are there tools describing what to cover when conducting a supervisory visit (e.g., a checklist)?				
If no to both 9.16 and 9.17, skip to question # 9.19					
9.18	If yes, are these guidelines and tools used by supervisors?				
		Yes	**No**	**NA**	**Comments**
9.19	Are supervisory visits conducted for staff at:				
a)	Regional				
b)	District				
c)	SDP				
If no, skip to question # 9.23					
9.20	If yes, what types of activities take place during the visits:				
a)	Review procedures for forecasting needs?				
b)	Review procedures for ordering products?				
c)	Observation of product storage?				
d)	Conducting physical inventory?				
e)	Review of logistics records and reports?				
f)	Discussion of budgeting for logistics activities?				
g)	Review of changes made since last supervisory visit?				
h)	On the job training to improve job performance?				
i)	Discussion of what is working and what is not?				
j)	Discussion of what help is needed? (staff, equipment, forms etc)				
9.21	Is there a documented schedule for supervision?				
If no, skip to question # 9.23					
9.22	If yes, are supervisory visits conducted according to the established schedule? How often do they take place? Are there any constraints to conducting supervisory visits?				

	Yes	No	NA	Comments
9.23 Are logistics staff periodically evaluated against job expectations (e.g. from their job description)?				
9.24 If a staff member's performance in logistics is not to standard, is the person provided with:				
a) In-service training?				
b) On-the-job training?				
c) Written instructions on how to improve?				
d) A coach or mentor?				

Staff Development in Logistics				
	Yes	No	NA	Comments
9.25 Does the program conduct periodic staff development activities (e.g., classroom training, coaching, on-the-job training, etc.)?				
	Yes	No	NA	Comments
9.26 Has training been given to current staff at all appropriate levels in the following areas:				
Completion and submission of LMIS reports?				
a) Proper storage of health products?				
b) Maintaining proper stock levels?				
c) Determining order quantities?				
d) Determining issue quantities?				
e) Estimating annual needs?				
f) Reviewing reports and records?				
g) Providing feedback and inputs?				

SECTION X- PRODUCT USE

Standard Treatment Guidelines and Universal Safety Precautions				
	Yes	No	NA	Comments
10.1 Do written standard treatment guidelines exist for conditions treated with commodities in the supply chain assessed?				
If no, skip to question # 10.4				
10.2 Specify the commodities in this supply chain that are required to comply with the standard treatment guidelines. Attach the list.				
	Yes	No	NA	Comments
10.3 Are they distributed to all the service delivery points?				
10.4 Are there written procedures for monitoring and supervising prescribing practices (e.g., monitoring number of products/drugs prescribed/dispensed per prescription)?				

E-498

If no, skip to question # 10.6					
10.5	Are they distributed to service providers at all levels?				
10.6	Do written universal safety precaution guidelines exist (e.g., disposing used needles, washing hands before and after contact with patient)?				
If no, skip to question # 10.8					
10.7	Are they distributed to service providers at all levels delivery points?				
10.8	What mechanisms and resources are in place to ensure the implementation of the standard treatment guidelines and universal safety precautions?				
a)	To what extent are they followed?				
b)	If not followed, what are the barriers to putting them into practice?				

		Yes	No	NA	Comments
10.9	Are commodities provided only to facilities that have staff trained, and equipped to use them (e.g., TB drugs only to DOT-trained facilities, IUDs only to sites with trained providers)?				
10.10					
a)	Are drug use studies conducted?				
b)	If so how often?				
c)	By whom?				

SECTION XI- FINANCE

Program Financing					
		Yes	No	NA	Comments
11.1	Does the program's budget include line items for *(specify the program):*				
Products?					
a)	Warehousing/storage?				
b)	Logistics management information system?				
c)	Transportation?				
d)	Logistics staff development?				
e)	Salaries for logistics staff?				
11.2	What is the program's annual budget: ♦ For drugs? _____ Reported year _____ ♦ For contraceptives? _____ Reported year _____ ♦ For logistics? _____ Reported year _____				
11.3	Who finances the program's annual budget? What percentage of the cost of products procured is locally financed?				
11.4	What is the process of developing the program's budget?				

11.5	What was the program's total annual expenditure for: *(Please also calculate the per capita expenditure)* ➢ Drugs? _____ Reported year _____ Per capita expenditure _____ ➢ Contraceptives? _____ Reported year _____ Per capita expenditure _____

		Yes	No	NA	Comments
11.6	Considering last available year's expenditure (capital and operating costs) is the budget sufficient? If not, why?				
11.7	Estimate the percentage of products bought from domestic vs. international suppliers.				

		Yes	No	NA	Comments
11.8	Are clients charged for:				
a)	Services?				
b)	Commodities?				
11.9	Are revenues generated from the cost recovery system used for:				
a)	Commodity costs?				
b)	Logistics costs?				
c)	Other costs?				
11.10	What approximate percentage of costs is recovered? If possible separate by commodity vs. logistics.				
11.11	Where is the money physically kept and managed? What is it used for?				

Donor Coordination					
		Yes	No	NA	Comments
11.12	Is there a process of coordination with donors for commodity supply?				
If no, skip to question #11.16					
11.13	Does this process occur at specified intervals?				
11.14	Describe the process and specify intervals.				
		Yes	No	NA	Comments
11.15	Does the program initiate the coordination with donors?				
11.16	Is there a mechanism or a unit that currently coordinates procurement and product shipment with donors?				
		Yes	No	NA	Comments
11.17	Are any products procured through a basket funding mechanism?				
If no, skip to #11.20					
11.18	Specify which products are procured through basket funding.				
11.19	Describe the process (e.g., timing, donors, etc).				
11.20	What are the program's future plans for local financing? Are there any plans by donors to phase out or reduce donations in the next 5 years?				

**DATA COLLECTION INSTRUMENT FOR EVALUATING
PROCESS AND IMPACT OF OPERATIONS RESEARCH**[1]

Project Number and Title:

Agreement Number:

Dates of Project:

Implementing Organization:

Collaborating Organization(s):

Project Summary:

Persons Interviewed:

P-1: Did the implementing/collaborating organization(s) actively participate in the design of the OR study?	1	2	3
P-2: Did the implementing/collaborating organization(s) actively participate in the implementation of the OR study?	1	2	3
P-3: Did the implementing/collaborating organization(s) participate in developing programmatic recommendations?	1	2	3
P-4: Did the study accomplish its research objectives?	1	2	3

---

[1] FRONTIERS/Tulane, 2001.

P-5: Was the intervention implemented as planned (or with some modifications)?	1	2	3

P-6: Was the study completed without delays (or other changes to the timeline) that compromised the validity of the research design?	1	2	3

P-7: Was continuity in key personnel maintained over the life of the OR project?	1	2	3

P-8: Was the study design methodologically sound (free of flaws that could have affected the final results)?	1	2	3

P-9: Was the research design feasible in the local context?	1	2	3

P-10: Did implementing/collaborating organizations judge the OR technical assistance to be useful and provided in a collegial manner?	1	2	3

P-11: Were results of the OR study judged to be credible/valid in the local context?	1	2	3

P-12: Was the research relevant for the national program?	1	2	3

P-13: Were the results disseminated to key audiences, including policy makers, program managers, and service providers?	1	2	3	

P-14: Are the results readily available in written form?	1	2	3	

I-1: Did the results indicate that the intervention was effective (i.e., that it improved service delivery in areas identified by the OR study)?	1	2	3	

I-2: Did the implementing/collaborating organization(s) "act on" the results (i.e., continue to implement the activities tested in the OR study after its completion if effective or not implement/discontinue this activity if ineffective)?	1	2	3	

I-3: If the intervention was effective and continued after the study, were the activities tested under the intervention still observable 36 months post-implementation?	1	2	3	NA

I-4: If the intervention was effective and continued after the study, was the intervention scaled up by the original implementing organization in the same country?	1	2	3	NA

I-5: If the intervention was effective and continued after the study, was the intervention adopted by another organization within the same country?	1	2	3	NA

I-6: If the intervention was effective and continued after the study, was the intervention replicated in another country?	1	2	3	NA

I-7: Was there a change in policy that can be linked to the OR project?	1	2	3

I-8: Did the implementing/collaborating organization(s) conduct subsequent OR studies?	1	2	3

I-9: Did the implementing/collaborating organization(s) conduct subsequent OR studies without external technical assistance?	1	2	3

I-10: Did the original donor fund new or expanded program activities based on the results of the OR study?	1	2	3

I-11: Did other donors provide new or expanded funding based on the results of the OR study?	1	2	3

C-1: Were there other factors (not mentioned above) that facilitated the conduct of the research project?

C-2: Were there other factors (not mentioned above) that facilitated the utilization of results from this operations research project?

C-3: Were there other factors (not mentioned above) that hindered the conduct of the research project?

C-4: Were there other factors (not mentioned above) that hindered the utilization of results from his operations research project?
C-5: Did the donor use the data from the OR study for a specific purpose?
C-6: Did the study include an assessment of the costs of the intervention?

Appendix G.1 Family Planning Services

Objectives and Priority Indicators		Tool
FP-I	**IMPROVE EFFECTIVENESS OF FAMILY PLANNING THROUGH QUALITY FAMILY PLANNING SERVICES**	
	ASSESSMENT	
1.1	% of facilities where at least 80% of the observed FP consultations included the following: ◄ new FP client consultations include assessment of essential components of reproductive and health information required to screen for method appropriateness ◄ pelvic examinations or method insertions met the all standards for quality (Aggregate indicator)	Observation
1.2	% of facilities where STI diagnosis and treatment is a part of the FP service	Inventory
1.3	Of facilities which use individual client cards, % where: ◄ during at least 80% of observed FP consultations, the Provider referred to the card either prior to or during the consultation ◄ for 100% of the observed FP consultations, the Provider wrote on the card	Inventory Observation
	COUNSELING	
1.4	% of facilities where at least 80% of the observed FP consultations included the following: ◄ clients were encouraged to express concerns or questions about methods ◄ clients who left with a method or referral for a method, received the specified key point on use and side effects ◄ a return visit was discussed with clinical method users ◄ some discussion occurred about risk of STIs and use of condoms for preventing STIs	Observation
1.5	% of facilities where at least 80% of interviewed clients ◄ who report a method problem also report discussing this with the provider ◄ who report a method problem also report discussing other methods with the provider ◄ report receiving information related to method use which includes possible side effects, what to do if they have problems, and when to return for follow-up ◄ report discussions which include all elements specified for good quality provider-client interaction ◄ respond correctly to a relevant question for their method ◄ report correctly if their method protects against STIs and AIDS	Exit Interview

---

FP-II	PROVIDE FAMILY PLANNING SERVICES UNDER CONDITIONS THAT SUPPORT QUALITY OF CARE	
	AVAILABILITY OF EQUIPMENT AND SUPPLIES	
2.1	% of facilities with all equipment and supplies for providing each clinical method of FP which is offered, including medications for treating STIs and blank individual client cards, if applicable, available (Aggregate indicator)	Inventory
	INFRASTRUCTURE AND TOOLS FOR QUALITY SERVICES	
2.2	% of facilities where FP services have: ◄ space routinely used for FP consultations that offers privacy ◄ space routinely used for FP examinations that offers privacy ◄ available elements for adequate infection prevention ◄ written protocols or guidelines for FP methods ◄ written protocols or guidelines for diagnosing and treating STIs ◄ visual aids for providing education to FP clients about FP methods ◄ visual aids for providing education to FP clients about STIs ◄ information booklet/pamphlets about FP which client can take home ◄ information booklet/pamphlets about STIs which client can take home	Inventory
2.3	% of facilities where, among observed consultations for FP clients ◄ at least 80% were conducted under conditions which assured both visual and auditory privacy ◄ at least 50% received instruction about family planning or related topics where visual aids or models were used	Observation
	INFORMATION SYSTEMS	
2.4	% of facilities where FP services: ◄ maintain up-to-date register which provides minimum standard (country specific) information on FP clients ◄ use individual FP client cards	Inventory
	MANAGEMENT	
2.5	% of facilities where, of interviewed Providers who provide FP services: ◄ at least 50% report having received training in issues related to FP within the prior 12 months ◄ at least 50% of services report having received at least one supervisory visit in the past six months	Provider Interview
FP-III	PROVIDE FAMILY PLANNING SERVICES UNDER A SYSTEM THAT PROMOTES UTILIZATION	
3.1	Average and median out-of-pocket expenditure for interviewed clients	Exit Interview
3.2	Compilation of % of clients who identify problems with service environment, by problem	Exit Interview
3.3	% of facilities where at least 80% of observed new FP clients were given verbal assurances of confidentiality	Observation
3.4	% of facilities where: ◄ Space routinely used for FP consultations offers privacy ◄ full range of clinical contraceptive methods as well as sterilization for males and females are provided	Inventory

Appendix G.2 Sexually Transmitted Infection (STI) Services

Objectives and Priority Indicators		Tool
STI-I	DECREASE HEALTH EFFECTS THROUGH QUALITY DIAGNOSIS AND TREATMENT SERVICES FOR STIS	
	ASSESSMENT	
1.1	% of facilities where at least 80% of the observed STI consultations included the following: ◄ elicited basic information required for diagnosis using Syndromic Approach ◄ used some form of laboratory examination (referral, taking specimen, or laboratory test actually conducted) as a part of diagnostic process ◄ included physical examination (external genitalia or pelvic examination) using appropriate methods for visualizing STI symptoms ◄ at least 80% of observed physical examinations met all quality standards (Aggregate indicator)	Observation
1.2	% of facilities where at least 80% of interviewed STI clients reported: ◄ providing specimen for laboratory examination or being prescribed a laboratory test ◄ being offered an HIV/AIDS test	Exit Interview
1.3	% of facilities: ◄ where antenatal clients are routinely offered testing for syphilis ◄ using clinical/etiologic diagnosis for STIs	Inventory
	TREATMENT	
1.4	% of facilities: ◄ where partner notification or follow-up is a routine part of the STI consultation system ◄ where ANC or FP clients with STIs can receive treatment through that service	Inventory
1.5	% of facilities where all observed STI clients who were prescribed medications: ◄ leave the facility with condoms ◄ leave the facility with all medications	Exit Interview
	COUNSELING	
1.6	% of facilities where at least 80% of observed consultations for FP included some discussion of STIs	Observation of FP Consultations
1.7	% of facilities where counseling for all observed STI clients who were prescribed medications included: ◄ discussion of partnership status ◄ encouragement to refer their partner(s) for treatment ◄ information on their diagnosis ◄ instruction on the importance of completing the full course of treatment ◄ a follow-up date to return for re-examination ◄ use of the condom for preventing transmission of STIs	Observation
1.8	% of facilities where: ◄ 100% of STI clients who report any laboratory or specimen exam, report they know why the test was ordered/specimen taken ◄ at least 80% of STI clients report that the Provider discussed prevention of STIs and HIV/AIDS and the client can mention at least one strategy ◄ at least 80% of STI clients report that the Provider discussed use of condoms to prevent STIs /HIV/AIDS	Exit Interview

Objectives and Priority Indicators	Tool
1.9 % of facilities where all interviewed clients who received prescriptions or medications report that: ◄ the Provider discussed use of condoms ◄ they were informed of their diagnosis ◄ they will take their medications until they are completed	Exit Interview
STI-II **PROVIDE STI SERVICES UNDER CONDITIONS THAT SUPPORT QUALITY OF CARE**	
AVAILABILITY OF EQUIPMENT AND SUPPLIES	
2.1 % of facilities where STI services have available: ◄ all equipment and supplies for providing STI consultation and examination using syndromic approach ◄ all equipment and supplies for providing STI consultation and examination using clinical/etiologic approach ◄ all essential medications for treating most common STIs ◄ have condoms at the service delivery site for STI consultations	Inventory
INFRASTRUCUTRE AND TOOLS FOR QUALITY SERVICES	
2.2 % of facilities where STI services have: ◄ space routinely used for STI consultations that offers privacy ◄ space routinely used for STI examinations that offers privacy ◄ elements for adequate infection prevention ◄ written protocols or guidelines for SYNDROMIC APPROACH to diagnosis and treatment of STIs ◄ written protocols or guidelines for CLINICAL/ETIOLOGIC diagnosis and treatment of STIs ◄ visual aids for providing education to STI clients about any aspect of STIs at STI service sites; at FP service sites ◄ information booklet/pamphlets about STIs which client can take home at STI service sites; at FP service sites	Inventory
2.3 % of facilities where, among observed consultations for STI clients: ◄ at least 80% were conducted under conditions which assured both visual and auditory privacy ◄ at least 80% of observed consultation and examination of STI clients were conducted using procedures for prevention of infection ◄ at least 50% receive education/counseling regarding STIs and related topics that used visual aids ◄ all observed STI clients who were prescribed medications received instructions on use of the condom for preventing transmission of STIs which included use of visual aids/model	Observation
INFORMATION SYSTEMS	
2.4 % of facilities where STI services maintain up-to-date register which provides (at minimum) information that STI was diagnosed, and type of STI, along with other minimum standard (country specific) information on STI clients	Inventory
MANAGEMENT	
2.5 % of facilities where, of interviewed Providers who provide STI services: ◄ at least 50% report having received continuing training on STI related issues within the prior 12 months ◄ at least 50% report having received at least one supervisory visit in the past six months	Provider Interview

	Objectives and Priority Indicators	Tool
STI-III	**PROVISION OF STI SERVICES UNDER A SYSTEM WHICH PROMOTE UTILIZATION**	
3.1	Average and median out-of-pocket expenditure for interviewed clients	Exit Interview
3.2	Compilation of % of clients who identify problems with service environment, by problem	Exit Interview
3.3	% of facilities providing: ◄ space routinely used for STI consultations that offers privacy ◄ space routinely used for STI examinations that offers privacy ◄ STI services through clinics which cater specifically to STI clients ◄ confidentiality protocols ◄ informed consent for testing for STIs	Inventory
3.4	% of facilities where: ◄ 80% of observed clients were told about confidentiality of the consultation ◄ all observed clients who had a specimen taken or a laboratory examination ordered were asked for their agreement or permission/the explanation included some explanation that it was to check for infection or a specific type of STI	Observation
STI-IV	**IMPROVE THE QUALITY OF LIFE AND MINIMIZE IMPACT OF OPPORTUNISTIC INFECTIONS FOR HIV/AIDS PATIENTS**	
	ASSESSMENT	
4.1	% of facilities which: ◄ provide laboratory diagnosis of HIV/AIDS and tuberculosis ◄ have laboratory ability to monitor effectiveness of HIV/AIDS therapy	Inventory
	TREATMENT	
4.2	% of facilities providing: ◄ medical management of opportunistic infections associated with HIV/AIDS positive clients ◄ anti-retroviral therapy ◄ programs to support palliative care of HIV/AIDS positive clients ◄ programs to provide social support for HIV/AIDS positive clients ◄ programs to provide service or referral for all essential aspects of HIV/AIDS treatment ◄ family planning counseling as a part of their services for HIV/AIDS clients	Inventory
STI-V	**PROVIDE HIV/AIDS SERVICES UNDER CONDITIONS THAT SUPPORT QUALITY OF CARE**	
	ESSENTIAL EQUIPMENT AND SUPPLIES	
5.1	% of facilities where HIV/AIDS services have available: ◄ essential laboratory equipment and supplies for diagnosing HIV/AIDS ◄ essential laboratory equipment and supplies for diagnosing tuberculosis ◄ anti-retroviral medications ◄ supply of essential medications for treating most common opportunistic infections for HIV/AIDS clients	Inventory

	INFRASTRUCTURE AND TOOLS FOR QUALITY SERVICES	
5.2	% of facilities where HIV/AIDS services have: ◄ space routinely used for HIV/AIDS consultations that offers privacy ◄ available elements for adequate infection prevention measures ◄ written protocols for clinical management of HIV/AIDS clients ◄ written protocols for referrals of HIV/AIDS clients ◄ booklets/pamphlets on HIV/AIDS available for clients to take home in each service area where HIV/AIDS; STI; FP services are provided ◄ written list of sources for referrals for HIV/AIDS clients	Inventory
	INFORMATION SYSTEMS	
5.3	% of facilities where HIV/AIDS services maintain: ◄ up-to-date register which provides minimum standard (country specific) information on HIV/AIDS clients ◄ records on numbers of clients receiving anti-retroviral therapy	Inventory
	MANAGEMENT	
5.4	% of facilities where HIV/AIDS services have a system for follow-up on referrals for HIV positive clients	Inventory
5.5	% of facilities where, of interviewed Providers who provide HIV/AIDS services: ◄ all either had training in HIV/AIDS during basic professional training or report having received continuing training on HIV/AIDS related issues within the prior 12 months ◄ where at least 50% report having received at least 1 supervisory visit in the past 6 months	Provider Interview
STI-VI	**PROVISION OF HIV/AIDS SERVICES UNDER A SYSTEM WHICH PROMOTE UTILIZATION**	
6.1	% of facilities which have/use: ◄ confidentiality protocols for HIV/AIDS ◄ informed consent for testing for HIV/AIDS ◄ HIV/AIDS counseling in space which offers privacy	Inventory
STI-VII	**PRO-ACTIVE ACTIVITIES TO DECREASE TRANSMISSION OF HIV/AIDS**	
7.1	% of facilities: ◄ offering pregnant women voluntary counseling and testing for HIV and education on MTC transmission ◄ providing VCT programs with related referral linkages	Inventory
7.2	% of facilities where at least 50% of interviewed Providers who offer <5 antenatal, newborn, STI, or HIV/AIDS services have been trained or received continuing education regarding MTC transmission.	Provider Interview

Appendix G.3 Maternal Health Services

	Objectives and Priority Indicators	Tool
MAT-I	**IMPROVE PREGNANCY OUTCOMES THROUGH QUALITY ANTENATAL CARE**	
	ASSESSMENT	
1.1	% facilities where, among observed ANC consultations: ◀ at least 75% of first visits included assessment of essential components of prior pregnancy history ◀ at least 75% were assessed for all major risk symptoms ◀ 100% had blood pressure taken ◀ at least 75% of clients 5+ months pregnant were asked about fetal movement and fetal heart rate was assessed, ◀ at least 75% of clients 8+ months pregnant had abdomen palpated ◀ at least 75% received good quality ANC procedures (aggregate of assessment indicators)	Observation Exit Interview
1.2	% of facilities using individual client cards, where: ◀ during at least 75% of observed ANC consultations, the Provider referred to the card either prior to or during the consultation. ◀ during 100% of the observed ANC consultations, the Provider wrote on the card	Inventory Observation
	TREATMENT	
1.3	% facilities where among observed ANC consultations, at least 75%: ◀ first visits had iron pills, tetanus toxoid, and anti-malarial medications discussed or prescribed ◀ clients who were prescribed iron tablets and/or anti-malarials leave the facility with the medications ◀ first visits were offered screening for HIV and/or VCT ◀ first visits were offered VDRL tests	Observation
1.4	% of facilities that: ◀ report routinely offering preventive and/or curative interventions for malaria; HIV/AIDS; STIs. ◀ provide STI treatment by ANC service providers	Inventory
1.5	% of facilities where at least 75% of interviewed ANC clients: ◀ report having been prescribed iron tablets and/or anti-malarial medications during any ANC visit	Exit Interview
	COUNSELING	
1.6	% of facilities where, of interviewed ANC clients: ◀ at least 75% report having been prescribed iron tablets and/or anti-malarial medications and having received an explanation about the medications during any ANC visit ◀ at least 75% report having received counseling during any ANC visit about nutrition during pregnancy ◀ at least 75% report having received counseling about exclusive breast feeding for at least six months during any visit ◀ 100% report having been counseled on warning signs or symptoms during pregnancy and can name at least one major symptom ◀ at least 75% report some pre-planning thought for delivery and/or discussion about delivery with the Provider	Exit Interview

MAT-II	PROVIDE ANTENATAL CARE UNDER CONDITIONS THAT SUPPORT QUALITY OF CARE	
	AVAILABILITY OF EQUIPMENT AND SUPPLIES	
2.1	% of facilities where ANC services: ◄ have available all equipment and supplies for providing basic ANC and postpartum examinations, including medications for treating STIs and blank individual client cards, if applicable ◄ have laboratory capacity for diagnostic tests for different risk conditions (malaria, syphilis, HIV/AIDS, anemia, proteinurea)	Inventory
	INFRASTRUCTURE AND TOOLS FOR QUALITY SERVICES	
2.2	% of facilities where ANC services: ◄ Space routinely used for ANC consultations offers privacy ◄ Space routinely used for ANC examinations offers privacy ◄ have available elements for adequate infection prevention ◄ have written protocols or guidelines for provision of ANC and management of problems during pregnancy ◄ have visual aids for providing education to ANC clients	Inventory
2.3	% of facilities where, among observed consultations for ANC clients: ◄ at least 75% were conducted under conditions which assured both visual and auditory privacy ◄ at least 50% received information about pregnancy or related topics where visual aids were used	Observation
	INFORMATION SYSTEMS	
2.4	% of facilities where ANC services: ◄ maintain up-to-date register which provides minimum standard (country specific) information on ANC clients ◄ monitor ANC coverage for a catchment population ◄ using individual ANC client cards	Inventory
	MANAGEMENT	
2.5	% of facilities where, of interviewed Providers who provide ANC services: ◄ at least 50% report having received training in issues related to maternity within the prior 12 months ◄ where at least 50% report having received at least one supervisory visit in the past 6 months	Provider Interview
MAT-III	PROVIDE ANC UNDER A SYSTEM THAT PROMOTES UTILIZATION	
3.1	Mean and median out-of-pocket expenditure for interviewed clients	Exit Interview
3.2	Compilation of % of clients who identify problems with service environment, by problem	Exit Interview
3.3	% of facilities where space routinely used for ANC consultations offers privacy	Inventory

MAT-IV	IMPROVE BIRTH OUTCOMES THROUGH PROVIDING DELIVERY CARE UNDER CONDITIONS WHICH SUPPORT QUALITY OF CARE	
	AVAILABILITY OF EQUIPMENT AND SUPPLIES	
4.1	% of facilities providing delivery services, where delivery services have available: ◄ all essential equipment and supplies, including medications for managing common complications of labor and delivery ◄ equipment for managing complications of miscarriage or abortion ◄ all essential equipment and supplies for managing emergency obstetric cases	Inventory
4.2	% of facilities which report providing cesarean sections and ◄ have all essential equipment and have skilled personnel to provide caesarian sections	Inventory
	INFRASTRUCTURE AND TOOLS FOR QUALITY SERVICES	
4.3	% of facilities where delivery services: ◄ have space routinely used for deliveries, offers privacy ◄ have available elements for adequate infection prevention ◄ have protocols or guidelines for managing normal deliveries and common complications ◄ use the partograph and have blank copies available	Inventory
	INFORMATION SYSTEMS	
4.4	% of facilities that: ◄ maintain up-to-date register which provides minimum standard (country specific) information on delivery clients ◄ monitor coverage for delivery care for a catchment population	Inventory
4.5	◄ % of facilities conducting cesarean sections that maintain up-to-date register indicating numbers and dates	Inventory
	MANAGEMENT	
4.6	% of facilities that have: ◄ a referral system for complicated maternity cases requiring a higher level of care ◄ systems for reviewing maternal deaths and "near miss" deaths	Inventory
4.7	In areas having high proportions of non-facility deliveries, % of facilities having active programs to increase proportion of safe deliveries	Inventory
4.8	% of facilities where, of interviewed Providers who provide delivery services ◄ at least 50% report receiving training in emergency obstetric care issues within the prior 12 months ◄ all interviewed Providers who conduct deliveries report having used a partograph within the prior month	Provider Interview
MAT-V	**PROVIDE DELIVERY SERVICES UNDER A SYSTEM THAT PROMOTES UTILIZATION**	
5.1	% of facilities offering 24-hour delivery services	Inventory
5.2	% of facilities where space routinely used for deliveries offers privacy	Inventory

MAT-VI	IMPROVE BIRTH OUTCOMES THROUGH PROVIDING NEWBORN CARE UNDER CONDITIONS WHICH SUPPORT QUALITY OF CARE	
	PREVENTIVE HEALTH SERVICES WHICH IMPACT ON NEWBORN	
6.1	% of facilities: ◀ providing basic ANC services, including specific components which impact on specific risk factors for newborn health (HIV, STI, Malaria, tetanus) ◀ having specified routine practices which promote a healthy newborn	Inventory
	AVAILABILITY OF EQUIPMENT AND SUPPLIES	
6.2	% of facilities: ◀ with basic equipment and supplies for managing fetal distress available ◀ with basic equipment for emergency care of newborn available	Inventory
	INFORMATION SYSTEMS	
6.3	% of facilities: ◀ which routinely weigh newborn ◀ with systems for reviewing newborn deaths and "near miss" deaths	Inventory
	MANAGEMENT	
6.4	% of facilities where, of interviewed Providers who provide newborn care: ◀ at least 50% report having been trained in issues related to neonatal care	Inventory

Appendix G.4 Child Health (CH) Services

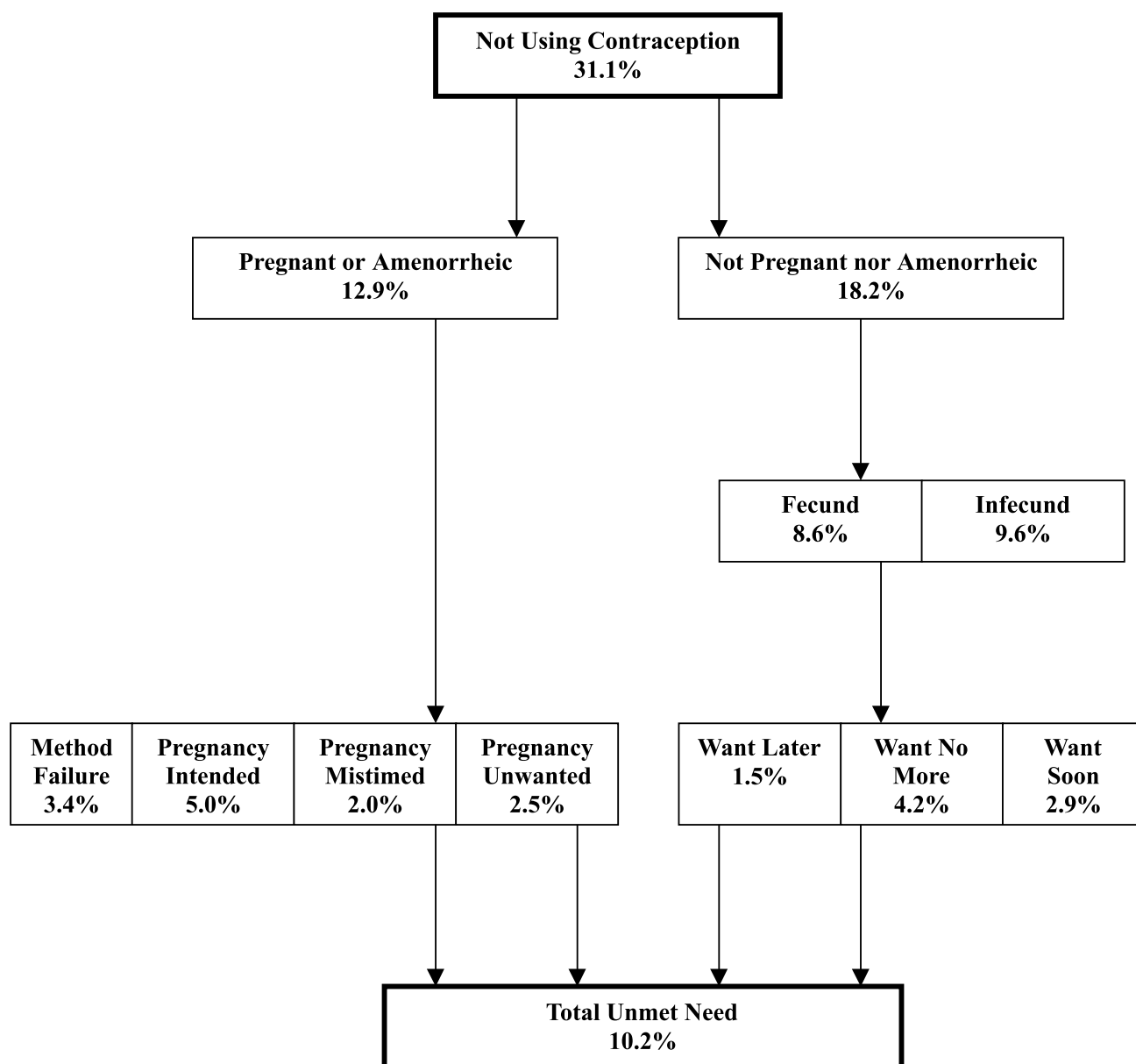|	Objectives and Priority Indicators	Tool
CH-I	**MINIMIZE MISSED OPPORTUNITIES FOR PREVENTIVE HEALTH INTERVENTIONS TO IMPROVE CHILD HEALTH**	
	NUTRITIONAL ASSESSMENT	
1.1	% of facilities where sick children are routinely ◀ weighed and weight plotted on growth chart ◀ assessed regarding immunization status	Inventory Observation
1.2	% of facilities where at least 80% of the observed consultations included discussion or consultation about ◀ weight/ growth of the child ◀ feeding/breast feeding practices for the child when not ill	Observation
1.3	% of facilities where at least 80% of the caretakers interviewed report that ◀ the sick child was weighed ◀ the child's weight or growth was discussed with them ◀ normal feeding/breast feeding practices when the child is not ill were discussed with them	Exit Interview
	COMPLETE MISSING IMMUNIZATIONS	
1.4	% of facilities offering immunization services on all days sick children are served	Inventory
1.5	% of facilities where at least 80% of the observed sick children who are below two years of age leave the facility with all needed vaccinations recorded on their cards	Exit Interview
CH-II	**IMPROVE HEALTH OUTCOME FOR ILL CHILD**	
	ASSESSMENT	
2.1	% of facilities where at least 80% of observed sick children : ◀ were assessed for presence of general danger signs ◀ were assessed for presence of cough or difficult breathing, diarrhea, and fever ◀ were assessed for feeding/breast feeding patterns during this illness ◀ had assessment which included all essential items covered in IMCI criteria (aggregate indicator)	Observation
2.2	% of facilities where routine components of sick child consultations include physical assessment (respiratory rate count, measured temperature, physical check for dehydration and/or anemia)	Inventory Observation
2.3	Of facilities which use individual child health cards, % where: ◀ during at least 80% of observed sick child consultations, the Provider referred to the card either prior to or during the consultation ◀ for 100% of the observed sick child consultations, the Provider wrote on the card	Inventory Observation
	TREATMENT	
2.4	% of facilities where at least 80% of observed sick children receive: ◀ appropriate treatment according to the classification of illness protocols for IMCI	Observation
2.5	% of facilities where giving the first dose of oral medication to the sick child prior to leaving the facility is a routine component of care	Inventory Observation

	COUNSELING	
2.6	% of facilities where at least 80% of caretakers of sick children were advised about: ◄ continued feeding or increasing amounts of food/breast-milk during this illness ◄ counseling covering all major issues recommended in IMCI guidelines ◄ how to give prescribed medications	Observation
2.7	% of facilities where at least 80% of caretakers of observed sick children: ◄ who are prescribed medications, leave facility with all prescribed medications ◄ report having received instructions on how to give oral medicines ◄ report feeling comfortable that they can provide the medications correctly ◄ report that the child was given a dose of at least one of the oral medications ◄ report receiving information on giving fluids and/or continued feeding for the child during this illness ◄ report being told about the illness and signs for which they should return to the facility	Exit Interview
2.8	% of facilities where visual aids were used during at least 50% of observed sick child consultations, when providing instruction/education to the caretaker	Observation
CH-III	**PROVIDE SICK CHILD CONSULTATION UNDER CONDITIONS THAT SUPPORT QUALITY OF CARE**	
	AVAILABILITY OF EQUIPMENT AND SUPPLIES	
3.1	% of facilities where sick child consultation services: ◄ have available all equipment and supplies for conducting examinations of sick children, including all essential medications, all essential supplies for immunization and, if applicable, blank child health cards	Inventory
	INFRASTRUCTURE AND TOOLS FOR QUALITY SERVICES	
3.2	% of facilities where sick child consultation services: ◄ have available elements for adequate infection prevention ◄ have written protocols or guidelines for management of childhood illnesses ◄ have visual aids for providing education to caretakers about child health	Inventory
	INFORMATION SYSTEMS	
3.3	% of facilities that: ◄ maintain up-to-date register which provides minimum standard (country specific) information on sick child consultations	Inventory
	MANAGEMENT	
3.4	% of facilities where, of interviewed Providers who provide care for sick children: ◄ at least one has received Integrated Management of Childhood Illness (IMCI) training at any time within the prior 59 months ◄ at least 50% report having received any training in specific health issues related to children within the prior 12 months ◄ at least 50% report having received at least one supervisory visit within the past six months	Provider Interview
CH-IV	**PROVIDE SICK CHILD CONSULTATIONS UNDER A SYSTEM THAT PROMOTES UTILIZATION**	
4.1	Average and median out-of-pocket expenditure for interviewed caretakers	Exit Interview
4.2	Compilation of % of interviewed caretakers who identify problems with service environment, by problem	Exit Interview

Ch-V	DECREASE CASES OF CHILDHOOD ILLNESS DUE TO IMMUNIZATION-PREVENTABLE DISEASES	
5.1	% of facilities which: ◀ provide both outreach and static vaccination services ◀ have appropriate storage system to maintain quality of vaccines ◀ have consistent supply of quality vaccines at immunization site ◀ have basic resources required to provide immunization services on service days ◀ monitor coverage rates for vaccinations for catchment population ◀ monitor drop-out rate for DPT ◀ have essential elements to provide immunization services under conditions which prevent infection	Inventory

Appendix G.5 Health Facility Services

Objective and Priority Indicators		Tool
FAC-I	**INCREASE AVAILABILITY AND FUNCTIONING LEVEL OF HEALTH SERVICES**	
	INFRASTRUCTURE	
1.1	% of facilities with: ◀ electricity which is available consistently when services are being provided ◀ functioning generator with fuel ◀ on-site water source	Inventory
	SERVICE AVAILABILITY	
1.2	% of facilities with: ◀ at least one secondary level staff assigned ◀ 24-hour emergency services and at least two secondary level staff assigned ◀ 24-hour access to emergency communication ◀ capacity to manage ill patients overnight ◀ at least some level of service for each of the following: well and sick child care, family planning, maternal health, and services for sexually transmitted illnesses ◀ curative services provided at least five days per week, and preventive child, maternal, and family planning services provided at least one day per week	Inventory
FAC-II	**OPERATE HEALTH SERVICES USING SYSTEMS WHICH PROMOTE QUALITY OF SERVICES**	
	MEDICINE AND SUPPLIES LOGISTICS	
2.1	% of facilities with: ◀ an up-to-date inventory for medicines, vaccines, and contraceptives ◀ reported routine system for ordering medications, vaccines, and contraceptives ◀ first-in first-out stock storage for medications and vaccines, to minimize drug expiration ◀ all medications (including vaccines) stored under conditions conducive to maintaining quality of items ◀ ___% of all essential drugs, 100% of offered contraceptive methods, 100% of offered vaccines, available the day of the facility visit. ◀ ___% of essential consumable supplies available the day of the facility visit ◀ no expired medications mixed with good medications. ◀ evidence of well-functioning logistic system for medications and consumable supplies (aggregate indicator with all of the above items)	Inventory
	EQUIPMENT MAINTENANCE	
2.2	% of facilities with: ◀ reported routine program for maintenance of major equipment ◀ more than ___ pieces of major equipment reported not functioning at the time of the facility visit ◀ reported routine system for maintaining or replacing small equipment ◀ more than ___ pieces (or %) of small equipment reported not functioning at the time of the facility visit	Inventory

	SYSTEMS FOR INFECTION PREVENTION	
2.3	% of facilities with: ◀ functioning equipment and reported systems for sterilizing/high level disinfection of equipment that meets standards ◀ safe disposal of sharp objects in all service provision areas where relevant ◀ decontamination solution in examination room, where relevant ◀ hand-washing soap and water available for all services surveyed during facility visit ◀ adequate disposal practices for hazardous waste ◀ evidence of well-functioning systems for prevention of infection (aggregate indicator with all of the above items)	Inventory
	MANAGEMENT	
2.4	% of facilities with: ◀ a formal system for reviewing administrative issues which meets routinely at least every six months ◀ a supervision visit from outside the facility at least once during the prior six months ◀ a known catchment population ◀ all interviewed secondary level staff having received continued education on a topic related to care of the sick child, maternal health, family planning, STIs, or HIV/AIDS at least once during the prior 12 months	Inventory Provider Interview
FAC-III	**PROVIDE HEALTH SERVICES UNDER CONDITIONS THAT PROMOTE UTILIZATION**	
	COMMUNITY COMPONENTS	
3.1	% of facilities with: ◀ community members as routine participants in the formal system for reviewing administrative issues for the facility ◀ participation in some type of formal plan which includes community/client financial support **plan will vary by country**	Inventory
	CLIENT COMFORT	
3.2	% of facilities with: ◀ protected client waiting area ◀ facilities with functioning latrine for client use ◀ maintenance at acceptable level of cleanliness	Inventory

ILLUSTRATION OF THE CALCULATION OF UNMET NEED FOR FAMILY PLANNING FROM THE PERU DHS (2000)

```
                    ┌─────────────────────────────┐
                    │  Not Using Contraception    │
                    │          31.1%              │
                    └─────────────────────────────┘
                       │                    │
         ┌─────────────────────────┐   ┌─────────────────────────────┐
         │ Pregnant or Amenorrheic │   │ Not Pregnant nor Amenorrheic│
         │         12.9%           │   │           18.2%             │
         └─────────────────────────┘   └─────────────────────────────┘
                                                    │
                                       ┌────────────────┬────────────────┐
                                       │   Fecund       │   Infecund     │
                                       │   8.6%         │   9.6%         │
                                       └────────────────┴────────────────┘
```

| Method Failure 3.4% | Pregnancy Intended 5.0% | Pregnancy Mistimed 2.0% | Pregnancy Unwanted 2.5% | | Want Later 1.5% | Want No More 4.2% | Want Soon 2.9% |

**Total Unmet Need
10.2%**

REFERENCES

Abdel-Tawab, N. and S. Hegazi. 2000. "Critical Analysis of Interventions against FGC in Egypt." Washington, DC: The Population Council, FRONTIERS.

AbouZahr, C. 1998. "Maternal Mortality Overview." In C.J.L. Murray and A.D. Lopez (eds) *Health Dimensions of Sex and Reproduction. Global Burden of Disease and Injury Series* III: 111-164. Geneva: World Health Organization.

AbouZahr, C. 1999. "Measuring Maternal Mortality: What Do We Need to Know?" In M. Berer and T.K.S. Ravindran (eds) *Safe Motherhood Initiatives: Critical Issues: Reproductive Health Matters: 13-23.* Oxford: Blackwell Science Ltd.

AbouZahr, C. and T. Wardlaw. Forthcoming 2001. "Maternal Mortality at the End of the Decade: What Signs of Progress?" *Bulletin of the World Health Organization.* Geneva: World Health Organization.

ACC/SCN. 1992. Second Report on the World Nutrition Situation. Geneva: ACC/SCN.

ACC/SCN News. 1994. *Maternal and Child Nutrition*. Number 11. Available online at: http://acc.unsystem.org/scn/publications/html/scnnews.html

ACC/SCN. 2000a. "Low Birthweight: Report of a Meeting in Dhaka, Bangladesh 14-17th June 1999." Eds. J. Pojda and L. Kelley. *Nutrition Policy Paper # 18.* Geneva: ACC/SCN in collaboration with ICDDRB

ACC/SCN, 2000b. "4th Report on the World Nutrition Situation: Nutrition throughout the Life Cycle." (in collaboration with the International Food Policy Research Institute). Available online at: http://acc.unsystem.org/scn/publications/rwns/4rwnscontents.html

Adair, L.A. 1998. "Testing Indicators for Use in Monitoring Interventions to Improve Women's Nutritional Status." *Working Paper 98-03*. MEASURE *Evaluation* Project. University of North Carolina, Chapel Hill: Carolina Population Center.

Akalin, M.Z., D. Maine, A. de Francisco, and R. Vaughan. 1997. "Why Perinatal Mortality Cannot Be a Proxy for Maternal Mortality." *Studies in Family Planning* 28: 330-335.

Adamchak, S., K. Bond, L. MacLaren, R. Magnani, K. Nelson, and J. Seltzer. 2000. *Monitoring and Evaluating Adolescent Reproductive Health Programs.* Washington, DC: FOCUS on Young Adults.

Akin J.S., D.K. Guilkey, P.L. Hutchinson, and M.T. McIntosh. 1998. "Price Elasticities of Demand for Curative Health Care with Control for Sample Selectivity on Endogenous Illness: An Analysis for Sri Lanka." *Health Economics* 7, 509-531.

Akweongo, P., S. Appiah-Yeboah, J.F. Phillips, E. Jackson, and E. Sakeah. 2001. "It's a Woman's Thing: Gender Roles Sustaining the Practice of Female Genital Mutilation among the Kassena-Nankana of Northern Ghana." Navrongo Health Research Centre, Ministry of Health, Box 114, Navrongo, Upper East Region, Ghana.

Anderson, M., J. Cosby, B. Swan, H. Moore, and M. Broekhoven. 1999. "The Use of Research in Local Health Service Agencies." *Social Science and Medicine* 49: 1007-1019.

Arnold, F. and A. Blanc. 1989. *Fertility Levels and Trend*s. Demographic and Health Surveys Comparative Studies No. 2. Columbia, Maryland: Institute for Resource Development/Macro International Inc.

Askov K, S. Legros, and V. Camacho. 2001. *Latin America and Caribbean Regional Initiative to Reduce Maternal Mortality*. LAMM General Briefer. Quality Assurance Project/ University Research Co. LLC and PAHO.

Babalola, S., C. Vondrasek, J. Brown, and R. Trao. 2001. "The Impact of a Regional Family Planning Promotion Initiative in West Africa: Evidence from Cameroon." *International Family Planning Perspectives* 27: 4.

Bailey, P., E. Bocalletti, G. Barrios, and Y. Cross. 2001. *Monitoring Changes in Maternal Health in Guatemala: Indicators to Measure Utilisation and Quality of Services.* Mothercare. (Unpublished)

Baker, J. M. Labbok, M. Lung'aho, and E. Sommerfelt. 1996. Tool Kit for Monitoring and Evaluating Breastfeeding Practices and Programs. Wellstart's Expanded Breastfeeding Promotion Program.

Baldridge, M. 2001. *Health Care Criteria for Performance Excellence*. Gaithersburg, MD: National Institute of Standards and Technology.

Bandura, A. 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bang, A.T., R.A. Bang, S.B. Baitule, M.H. Reddy, and M.D. Deshmukh. 1999. "Effect of Home-based Neonatal Care and Management of Sepsis on Neonatal Mortality: Field Trial in Rural India (comments)." *Lancet* 354, 9194: 1955-1961.

Banwell, S., M. Ellsberg, and V. Frye. 2001. Personal communication.

BEMFAM and Macro International. 1997. *Pesquisa Nacional sobre Demografia e Saúde,1996.* Brazil: Sociedade Civil Bem-estar Familiar and Calverton, MD: Macro International Inc.

Berer, M. and T.K. Ravindran. 2001. "Preventing Maternal Mortality: Evidence, Resources, Leadership Action." *Safe Motherhood Initiatives, Critical Issues*: 3-9. Oxford: Blackwell Science Ltd.

Berg, C., I. Danel, and G. Mora. 1996. *Guidelines for Maternal Mortality Epidemiological Surveillance.*Washington, DC: OPS.

Bertrand, J.T. and L. Brown. 1997. *The Working Group on the Evaluation of Family Planning Operations Research. Final Report.* The EVALUATION Project. University of North Carolina, Chapel Hill: Carolina Population Center.

Bertrand, J.T. and D.L. Kincaid. 1996. *Evaluating Information-Education-Communication (IEC) Programs for Family Planning and Reproductive Health.* WG-IEC-03. University of North Carolina: Carolina Population Center. Available online at http:/www.cpc.unc.edu/measure.

Bertrand, J.T. and C. Marin. 2001. *Operations Research: Measuring its Impact on Service Delivery and Policy.* Washington, DC: The Population Council, FRONTIERS. (Unpublished)

Bertrand, J.T., R. Magnani, and J. Knowles. 1994. *Handbook of Indicators for Family Planning Program*

Evaluation. The EVALUATION Project. University of North Carolina, Chapel Hill: Carolina Population Center.

Bertrand, J.T., R. Magnani, and N. Rutenberg. 1996. *Evaluating Family Planning Programs with Adaptations for Reproductive Health*: The EVALUATION Project. University of North Carolina, Chapel Hill: Carolina Population Center.

Bertrand, J.T. and A. Tsui (eds). 1995. *Indicators for Reproductive Health Program Evaluation*. Chapel Hill, North Carolina, University of North Carolina at Chapel Hill: Carolina Population Center, EVALUATION Project.

Bertrand, J.T., K. Hardee, R. Magnani, and M. Angle. 1995. "Access, Quality of Care and Medical Barriers In Family Planning Programs." *International Family Planning Perspectives* 21, 2: 64-69,74.

Bessinger, R. and J.T. Bertrand. 2001. "Monitoring Quality of Care in Family Planning Programs: A Comparison of Observation and Client Exit Interviews." *International Family Planning Perspectives.* 27(2):63-70.

Blanc, A.K. 2001. "The Effect of Power in Sexual Relationships on Sexual and Reproductive Health: an Examination of the Evidence." *Studies in Family Planning*. 32[3]: 189–213

Blanc, A., S. Curtis, and T. Croft. 1999. "Does Contraceptive Discontinuation Matter? Quality of Care and Fertility Consequences." MEASURE *Evaluation* Project Working Paper Series WP-99-14. University of North Carolina, Chapel Hill: Carolina Population Center.

Boerma, J.T., K.I. Weinstein, S.O. Rutstein, and A.E. Sommerfield. 1996. "Data on Birthweight in Developing Countries: Can Surveys Help?" *Bulletin of the World Health Organization* 74, 2: 209-216.

Bongaarts, J. 1990. "The Measurement of Unwanted Fertility." *Population and Development Review* 16, 3: 487-506.

Bongaarts, J. 1991. "The KAP–Gap and the Unmet Need for Contraception." *Population and Development Review* 17, 2: 293–313.

Boulier, B. 1985. "The Effects of Family Planning Programs on Fertility in the Developing World." *World Bank Staff Papers* 677.

Boyer, D. and D. Fine. 1992. "Sexual Abuse as a Factor in Adolescent Pregnancy and Child Maltreatment." *Family Planning Perspectives* 24, 1: 4-11,19.

Bratt, J.H., J. Foriet, P. Chein, C. West, B. Janowitz, and T. De Vargas. 1999. "A Comparison of Four Methods for Measuring Clinician Time Use." *Health Policy and Planning* 14, 4: 374-381.

Brown, L., A. Lafond, and K. Macintyre. 2001. *Measuring Capacity Building. Technical Report.* MEASURE *Evaluation* Project. University of North Carolina, Chapel Hill: Carolina Population Center.

Bruce, J. 1990. "Fundamental Elements of Quality of Care: A Simple Framework." *Studies in Family Planning* 21, 2: 61-91.

Buchan, J. 2001. "Human Resource Quality and Performance: What now, What next?" A presentation given at USAID/Washington.

Bulatao, R. and J. Ross. 2000. "Rating Maternal and Neonatal Health Programs in Developing Countries." MEASURE *Evaluation* Working Paper Series WP-00-26. University of North Carolina, Chapel Hill: Carolina Population Center.

Bulatao, R. and J. Ross. 2001. "Do Health Services Reduce Maternal Mortality? Evidence from Ratings of Maternal Health Programs?" MEASURE *Evaluation* Working Paper Series WP-01-39. University of North Carolina, Chapel Hill: Carolina Population Center.

Burt, Martha. 1980. "The Cultural Myths and Support for Rape." *Journal of Personality and Social Psychology* 38, 2: 217-230.

Busza, J. and L. Lush. 1999. Table from: "Planning Reproductive Health in Conflict: A Conceptual Framework." *Social Science and Medicine* 49: 155-171.

Caldwell, J.C., I.O. Orubuloye, P. Caldwell. 1999. "Obstacles to Behavioural Change to Lessen the Risk of HIV Infection in the African AIDS Epidemic: Nigerian Research." In: Caldwell J., P. Caldwell, J. Anarfi, K. Awusabo-Asare, J. Ntozi, I.O. Orubuloye, J. Marck, W. Crosford, R. Colombo and E. Hollings. *Resistances to Behavioural Change to Reduce HIV/AIDS Infection in Predominantly Heterosexual Epidemics in Third World Countries*. Canberra, Australia: Australian National University, National Centre for Epidemiology and Population Health, Health Transition Centre.

Caldwell, J.C., I.O. Orubuloye, and P. Caldwell. 2000. "Female Genital Mutilation: Conditions of Decline." *Population Research and Policy Review* 19: 233-54.

Campbell, O. 1999. "Measuring Progress in Safe Motherhood Programmes: Uses and Limitations of Health Outcome Indicators." In M. Berer and T.K.S. Ravindran (eds*) Safe Motherhood Initiatives: Critical Issues: Reproductive Health Matters 31-42.* Oxford: Blackwell Science Ltd.

Campbell, O., V. Filippi, M. Koblinsky, T. Marshall, J. Mortimer, R. Pittrof, C. Ronsmans, and L. Williams. 1997. *Lessons Learnt: A Decade of Measuring the Effects of Safe Motherhood Programs.* London, England: Maternal and Child Epidemiology Unit. London School of Hygiene and Tropical Medicine.

Campbell, O.M.R. and W.J. Graham. 1990. *Measuring Maternal Mortality and Morbidity: Levels and Trends.* London, England: Maternal and Child Epidemiology Unit. Publication No 2. London School of Hygiene and Tropical Medicine.

Campbell, O.M.R. and W.J. Graham. 1991. *Measuring the Determinants of Maternal Morbidity and Mortality: Defining and Selecting Outcomes and Determinants, and Demonstrating Associations.* London, England: Maternal and Child Epidemiology Unit. Publication No 4. London School of Hygiene and Tropical Medicine.

The Center for Reproductive Law and Policy. 2001. *Female Circumcision/Female Genital Mutilation: Global Laws and Policies Towards Elimination.* Factsheet. Available online at: http://www/crlp.org/pub_fac_fgmicpd.html

The Center for Reproductive Law and Policy. 1999. *Nations Worldwide Support a Woman's Right to Choose Abortion*. Factsheet. Available online at http://www.crlp.org/pub_fac_atkwwsup.html

Chayovan, N., A. I. Hermalin, and J. Knodel. 1984. "Measuring Accessibility to Family Planning Services in Thailand." *Studies in Family Planning* 15, 5:201–211.

Child Health Research Project. 1999. *Reducing Perinatal and Neonatal Mortality. Report of a Meeting in Baltimore, Maryland May 10-12.* Child Health Research Project Special Report. 3, 1.

Christian, P., K.P. West Jr., S.K. Khatry, J. Katz, S.R. Shrestha, E.K. Pradhan, S.C. LeClerq, and R.P. Pokhrel. 1998a. "Night Blindness of Pregnancy in Rural Nepal-Nutritional and Health Risks." *International Journal of Epidemiology* 27: 231-237.

Christian, P., K.P, West Jr., S.K. Khatry, J. Katz, S.R. Shrestha, E.K. Pradhan, S.C. LeClerq, E.K. Pradhan, and S.R. Shrestha. 1998b. "Vitamin A or B-carotene Supplementation Reduces but Does Not Eliminate Maternal Night Blindness in Nepal." *Journal of Nutrition* 128:1458-1463.

Christian, P., K.P. West Jr., S.K. Khatry, E. Kimbrough-Pradhan, S.C. LeClerq, J. Katz, S.R. Shrestha, S.M. Dali, and A. Somme. 2000. "Night Blindness during Pregnancy and Subsequent Mortality among Women in Nepal: Effects of Vitamin A or B-carotene Supplementation." *American Journal of Epidemiology*. 152: 542-547.

Clark, W.H. 1998. *Gender and Social Context in Population Council Operations Research: Lessons From the Past, Thoughts for the Future.* New York, NY: The Population Council.

Cross, H., N. Jewell, and K. Hardee. 2001. "Addressing Operational Policies through Policy Reform: Improving Reproductive Health Programs." POLICY Working Paper. Washington, DC: The Futures Group International.

DELIVER/JSI. 2002. Logistics Indicator Assessment Tool. John Snow, Inc.: Arlington, VA.

DELIVER/JSI. 2002. Logistics Systems Assessment Tool. John Snow, Inc.: Arlington, VA.

Dietrich J., D. Guilkey, and D. Mancini. 1998. "Family Planning, Maternal / Child Health,and Sexually-Transmitted Diseases in Tanzania: Multivariate Results Using Data from the 1996 Demographic and Health Survey and Service Availability Survey." MEASURE *Evaluation* Project, WP-98-01USAID. Chapel Hill, North Carolina, University of North Carolina at Chapel Hill: Carolina Population Center.

Dixon–Mueller, R. and A. Germain. 1992. "Stalking the Elusive 'Unmet Need' for Family Planning." *Studies in Family Planning* 23, 5:330–335.

Dobash, R.P., R.E. Dobash, M. Wilson, and M. Daly. 1992. "The Myth of Sexual Symmetry in Marital Violence." *Social Problems* 39, 1: 71-91.

Douglas, M. 1985. *Risk Acceptability According to the Social Sciences*. New York, NY: Russell Sage Foundation.

Eggleston, E., J. Jackson, W. Rountree, and Z. Pan. 2000. "Evaluation of a Sexuality Education Program for Young Adolescents in Jamaica." *Pan American Journal of Public Health* 7, 2: 102-112.

El-Zanaty, F., E.M. Hussein, G.A. Shawky, A.A. Way, and S. Kishor. 1996. *Egypt Demographic and Health Survey 1995.* Cairo, Egypt: National Population Council and Calverton, MD: Macro International Inc.

El-Zantay, F. and A. Way. 2001. *Egypt Demographic and Health Survey, 2000.* Calverton, Maryland [USA], Ministry of Health and Population [Egypt], National Population Council and ORC Macro.

Espeut, D. 1998. "Perinatal Mortality in Developing Countries: a Review of the Current Literature and Methodological Issues in Community-based Assessment." *MotherCare Technical Working Paper No. 5.* Arlington, VA: John Snow Inc.

The EVALUATION Project. 1998. *The Central American HIV/AIDS Prevention Activity Evaluation Plan.* New Orleans, LA: Tulane University.

Filippi, V., C. Ronsmans, T. Gandaho, W. Graham, E. Alihonou, and P. Santos. 2000. "Women's Reports of Severe (near-miss) Obstetric Complications in Benin." *Studies in Family Planning* 31, 4: 309-324.

Fishbein, M. and I. Ajzen. 1975. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research.* Reading, MA: Addison-Wesley.

Fisher, A., J. Laing, J. Stoeckel, and J. Townsend. 1991. *Handbook for Family Planning Operations Research Design.* New York, NY: The Population Council.

Foreit, J.R., M.E. Gorosh, D.G. Gillespie, and C.G. Merritt. 1978. "Community–based and Commercial Contraceptive Distribution: An Inventory and Appraisal." *Population Report*s Series J, 19, March: J1-J29.

Foreit, K. 1992. "Unmet Demand for Contraception vs. Unmet Demand for Appropriate Contraception." Paper Presented at the 120th Annual Meeting of the American Public Health Association. Washington, DC.

Foreman, M. (ed). 1999. *AIDS and Men: Taking Risks or Taking Responsibility*. London, England: Panos Institute.

Fortney, J. A. and J.B. Smith. 1999. "Measuring Maternal Morbidity." In M. Berer & T. K. Sundari (eds), *Safe Motherhood Initiatives: Critical Issues:* 43-52. Oxford: Blackwell Science Ltd.

FPLM/John Snow, Inc., 2000. *The Logistics Handbook: A Practical Guide for Supply Chain Managers in Family Planning and Health Programs.* Arlington, VA: Family Planning Logistics Management/ John Snow, Inc., for the U.S. Agency for International Development (USAID).

FPMD, 1993. "Using CQI to Strengthen Family Planning Programs." *The Family Planning Manager* 2, No. 1. Boston, MA: Management Sciences for Health.

FPMD. 1999. "Human Resources: Managing and Developing your Most Important Asset." *The Manager* 8, No. 1. Boston: MSH.

FRONTIERS/Tulane. 2001. Data Collection Instrument for Evaluating Process and Impact of Operations Research. Washington, DC: The Population Council.

Frye, V., S. Banwell, and M. Ellsberg. 2001. *Reproductive Health Program Responses to Gender-Based Violence Against Women: Conceptualizing Indicators on Monitoring and Evaluation.*(Unpublished)

Gloyd, S., S.Chai, and M.A. Mercer. 2001. "Antenatal Syphilis in Sub-Saharan Africa: Missed Opportunities for Mortality Reduction." *Health Policy Planning* 16, 1: 29-34.

Graham, W. 2001. "Monitoring and Evaluating: The Impact of Skilled Attendance." In J. Fullerton (ed). *Skilled Attendant at Delivery: A Review of the Evidence.* New York, NY: Interagency Group for Safe Motherhood/ Family Care International.

Graham, W. and P. Airey. 1987. "Measuring Maternal Mortality: Sense and Sensitivity.*" Health Policy and Planning* 2, 4: 323-333.

Graham, W., W. Brass, and R.W. Snow. 1989. "Estimating Maternal Mortality: The Sisterhood Method." *Studies in Family Planning* 20, 3: 125-135.

Graham, W. and V. Filippi. 1994. *Monitoring Maternal Health Goals: How Well Do the Indicators Perform?* London, England: Maternal and Child Epidemiology Unit, London School of Hygiene and Tropical Medicine.

Graham, W., V. Filippi, and C. Ronsmans. 1996. "Demonstrating Programme Impact Using Maternal Mortality." *Health Policy and Planning* 11, 1: 16-20.

Graham, W. and J. Bell. 2000. "Monitoring and Evaluating Skilled Attendance at Delivery: Trials and Tribulations." Paper presented at the Technical Consultation "Ensure Skilled Attendance at Delivery April 25-27." Geneva: World Health Organization.

Graham, W., P. Wagaarachchi, G. Penney, A. McCaw-Binns, K. Yeboah Anturi, and M.H. Hall. 2000. "Criteria for Clinical Audit of the Quality of Hospital-based Obstetric Care in Developing Countries." *Bulletin of the World Health Organization* 78, 5: 614-620.

Greene, M.E., 1999. "The Benefits of Involving Men in Reproductive Health." Paper presented in November 1999 at meeting of the Association for Women in Development, USAID and APHA annual meeting.

Gupta, G.R. 2000. "Gender, Sexuality, and HIV/AIDS: The What, the Why and the How." Plenary address XIIIth International AIDS Conference. Durban, South Africa.

Gutierrez, M. 2001. "Indicators for Assessing CGI Efforts." The Manager's Electronic Resource Center. Available online at: http://erc.msh.org/mainpage.cfm?file=2.2.6htm&module=quality&language=English

Haigh, S. and J. Megarity. 1998. Measuring Web Site Usage: Log File Analysis.

Heise, L., M. Ellsberg, and M. Gottemoeller. 1999. "Ending Violence against Women." *Population Reports. Series L: Issues in World Health* 11: 1-43.

Heise, L., L. Moore, and N. Toubia. 1995. *Sexual Coercion and Reproductive Health: A Focus on Research.* New York, NY: The Population Council.

Helzner, J.F. 1996. "Men's Involvement in Family Planning." *Reproductive Health Matters* No.7, May: 146-153.

Hermalin, A.I. and B. Entwisle. 1985. "Future Directions in the Analysis of Contraceptive Availability." Proceedings from the International Population Conference, June 5-12, Florence, Italy.

Hermalin, A.I. and B. Entwisle. 1988. "The Availability and Accessibility of Contraceptive Services." In R.J. Lapham and W.P. Mauldin (eds) *Organizing for Effective Family Planning Program*s. Washington, DC: National Academy Press.

Huezo C. and U. Malhotra. 1993 *Choice and Use-Continuation of Methods of Contraception: a Multicentre Study.* London, England: International Planned Parenthood Federation. 3, 163.

Hutchings, J., G. Perkin, and L. Saunders. 1987. "The Effect of Contraceptive Technology on the Program Environment." In R.J. Lapham and G.B. Simmons (eds) *Organizing for Effective Family Planning Programs.* Washington, DC: National Academy Press.

Institute National de la Statistique et de la Demographie, and Macro International Inc. 2000. Enquete Demographique et de Sante, Burkina Faso 1998-1999. Calverton, Maryland (USA): Macro International Inc.

Institute of Medicine/National Academy of Sciences. 1990. *Nutrition During Pregnancy*. Washington, D.C.: National Academy Press.

Interagency Gender Working Group. 2000. *Guide for Incorporating Gender Considerations in USAID's Family Planning and Reproductive Health RFAs and RFPs.* Washington D.C.: USAID. RFA Subgroup. Program Implementation Subcommittee.

Inter-Agency Group for Safe Motherhood. 2000. *Implementing the Safe Motherhood Action Agenda: A Resource Guide.* Available online at http://www.safemotherhood.org/

International Planned Parenthood Federation/Western Hemisphere Region Inc. 2000a. *Self-Assessment Module on Integrating SIT/HIV/AIDS Services into Sexual and Reproductive Health Programs.* New York, NY: IPPF/WHR.

International Planned Parenthood Federation/Western Hemisphere Region Inc. 2000b. *Manual to Evaluate Quality of Care from a Gender Perspective*. New York, NY: IPPF/WHR.

Ipas. 2000. Maternal Health: Implications for Children and Adolescents. Factsheet. Available online at: www.ipas.org/arch/index.htm.

Ipas/PRIME, 1998. "Post Abortion Care Services: an Update from PRIME." Resources for Women's Health. Volume 1, Number 2. Carrboro, North Carolina: Ipas.

Iskandar, M. and S. Indrawati. 1996. *Indonesia: Utilization of Completed Operations Research Studies. Final report.* Jakarta: The Population Council.

IVACG/WHO/UNICEF. 1998. "Guidelines For The Use Of Iron Supplements To Prevent And Treat Iron Deficiency Anemia." Geneva: WHO.

Izett S. and N. Toubia. 1999. "Learning about Social Change: A Research and Evaluation Guidebook Using Female Circumcision as a Case Study." New York, NY: Research Action and Information Network for Bodily Integrity of Women.

Jacobson, J. 2000. "Transforming Family Planning Programmes: Towards a Framework for Advancing the Reproductive Rights Agenda." *Reproductive Health Matters:* 8, 15.

Jain, A. 1995. "Implementing the ICPD's Message." *Studies in Family Planning* 26, 5: 296-298.

Jain, A. 2001. "Implications for Evaluating the Impact of Family Planning Programs with a Reproductive Health Orientation." *Studies in Family Planning* 32, 3:220-229.

James, W.P.T., et al. 1988. "Definition of Chronic Energy Deficiency in Adults: Report of a Working Party of the International Dietary Energy Consultative Group." *European Journal of Clinical Nutrition* 42: 969-981.

Janowitz, B. and J.H. Bratt. 1994. *Methods for Costing Family Planning Services.* UNFPA and Research Triangle Park, NC: Family Health International.

Kenney, G.M. 1993. "Assessing Legal and Regulatory Reform in Family Planning." *OPTIONS for Population Policy.* Policy Paper Series No. 1. Washington, DC.: The Futures Group International.

Kidman, C. 1993. "Non-Consensual Sexual Experience and HIV Education: An Educator's View." *SIECUS Report* 21, 4: 9-12.

Kim, Y.M., J. Rimon, K. Winward, C. Corso, Mako IV, S. Lawal, S. Babalola, and D. Huntington. 1992. "Improving the Quality of Service Delivery in Nigeria." *Studies in Family Planning* 23, 2: 118-27.

Kim, Y.M. and C. Lettenmaier. 1995. *Tools to Assess Family Planning Counseling: Observations and Interview*. IEC Research Tools. Baltimore, MD: John Hopkins Center for Communication Programs.

Kim, Y.M., F. Putjuk, E. Basuki, and A. Kols. 2000. "Self-Assessment and Peer Review: Improving Indonesian Service Providers' Communication with Clients." *International Family Planning Perspectives* 26, 1: 4-12.

Kincaid, D L., 2000. "Mass Media, Ideation, and Behavior: A Longitudinal Analysis of Contraceptive Change in the Philippines." *Communication Research* 27, 6: 723-763.

Kincaid, D. L., P.L. Coleman, and J.G. Rimon. 1995. "Image and the Adoption of Contraceptive Behavior: Impact of the Philippine National Communication Campaign in 1993." Paper presented at the annual meeting of the International Communication Association May 25-29. Albuquerque, NM.

Kincaid D.L., M.E., Figueroa, D. Storey, and C. Underwood. 1999. "Ideation and Contraceptive Behavior: The Relationship Observed in Five Countries" Paper presented at the annual meeting of the Population Association of America. March 25-27. New York, NY.

Kirby, D. 2001. *Emerging Answers: Research Findings on Programs to Reduce Teen Pregnancy.* Washington, DC: The National Campaign to Prevent Teen Pregnancy.

Kirkpatrick, D. 1998. *Evaluating Training Programs: The Four Levels*, 2nd ed. San Francisco, CA: Berrett-Koehler Publishers, Inc.

Kishor, S. 1999. *A Framework for Understanding the Role of Gender and Women's Status in Health and Population Outcomes.* MEASURE DHS[+] (Unpublished)

Koblinsky, M., K. Mc Laurin, P. Russell-Brown, and P. Gorbach (eds). 1995. *Indicators for Reproductive Health Program Evaluation.* Final Report of the Subcommittee on Safe Pregnancy. Chapel Hill, NC: The EVALUATION Project.

Koblinsky, M., C. Conroy, N. Kureshy, and C. Staton. 2000. *Issues in Programming for Safe Motherhood.* MotherCare. Arlington, VA: John Snow Inc..

Krasovec, K and M.A. Anderson. 1991a. "Maternal Nutrition and Pregnancy Outcomes: Anthropometric Assessment." *Scientific Publication No. 529* Washington DC: Pan American Health Organization.

Krasovec, K and M.A. Anderson. 1991b. "Maternal Anthropometry for Prediction of Pregnancy Outcomes: Memorandum from a USAID/WHO/PAHO/MotherCare meeting." *Bulletin Of The World Health Organization.* 1991;69(5):523-32.

Kumaranayake, L., J. Pepperall, H. Goodman, and A. Mills. 1998. *Costing Guidelines for HIV/AIDS Prevention Strategies (A companion volume to COST ANALYSIS IN PRIMARY HEALTH CARE: A Training Manual for Programme Managers* [WHO]*)*. London, England: Health Economics and Financing Programme, London School of Hygiene and Tropical Medicine.

Laboratoire de Santé Communautaire du Bazega. 1998a. *Evaluation de la prevalence, de la typologie et des complications liées à l'excision chez les patientes frequentant les formations sanitaires du Bazega.* Serie Documentaire no. 21 Ministère de la Santé, Unité d'Enseignement de la Recherche en Démographie (UERD), and The Population Council.

Laboratoire de Sante Communautaire du Bazega. 1998b. *Etude participative pour l'identification des stratégies communautaires de lutte contre la pratique le l'excision dans le Bazega.* Serie Documentaire 22. Ministère de la Santé, Unité d'Enseignement de la Recherche en Démographie (UERD), and The Population Council.

The Lancet Editorial. April 10, 1993. "Reproductive Freedom for Refugees." *The Lancet* 341: 929-30.

The Lancet. 2000. "Caesarean Section on the Rise." *The Lancet* 356, 9243: 1697.

Lapham, R.J. and W.P. Mauldin. 1984. "Family Planning Program Effort and Birthrate Decline in Developing Countries." *International Family Planning Perspectives* 10, 4:109–118.

Leonard, A.H. and J. Winkler. 1991. *A Quality of Care Framework for Abortion Care. Advances in Abortion Care* 1, 1. Carrboro, NC: Ipas.

Lightbourne, R.E. 1985. "Desired Number of Births and Prospects for Fertility Decline in 40 Countries." *International Family Planning Perspectives* 11, 2: 34–47.

Lightbourne, R.E. 1987. "Reproductive Preferences and Behavior." In J. Cleland and C. Scott (eds) *The World Fertility Survey: An Assessmen*t. London, England: Oxford University Press.

Lung'aho, M.S., J. Baker, M. Labbock, M. Lung'aho, and E. Sommerfelt. 1996. *Tool Kit for Monitoring and Evaluating Breastfeeding Practices and Programs.* Wellstart International and USAID.

Lutz, W. 1990. *Distributional Aspects of Human Fertility: A Global Comparative Stud*y. New York, NY: Academic Press.

Maine D., J. McCarthy, and V.M. Ward. 1992. *Guidelines for Monitoring Progress in the Reduction of Maternal Mortality: A Work in Progress.* New York, NY: UNICEF.

Maine D, T.M. Wardlaw, V.M. Ward, J. McCarthy, A. Birnbaum, M.Z. Akalin , and J.E. Brown. 1997. *Guidelines for Monitoring the Availability and Use of Obstetric Services.* 2nd Edition. New York, NY: UNICEF.

Maiuro R.D., R.R. Vitaliano, N.K. Sugg, D.C. Thompson, F.P Rivara, and R.S. Thompson. 2000. "Development of a Health Care Provider Survey for Domestic Violence: Psychometric Properties." *American Journal of Preventive Medicine* 19, 4: 245-252.

Mauldin, W.P. and J.A. Ross. 1991. "Family Planning Programs: Efforts and Results, 1982–89." *Studies in Family Planning* 22, 6: 350–367.

Mauldin, W.P., J.A. Ross, J. Kekovole, Barkat-e-Khuda, and A. Barkat. 1995. "Direct and Judgmental Measures of Family Planning Program Inputs." *Studies in Family Planning* 26, 5: 287-295.

McCaffery, J., M. Louma, C. Newman, S. Rudy, A. Fort, and F. Rosensweig. 1999. *Performance Improvement: Stages, Steps, and Tools.* Chapel Hill, NC: Intrah.

McCarthy, B., J. Lawn, and S. Rae Ross. 2001. *The Healthy Newborn: A Reference Manual for Program Managers.* CARE, Center for Disease Control, CCHI, and World Health Organization. (Unpublished)

McCarthy, J. and D. Maine. 1992. "A Framework for Analyzing the Determinants of Maternal Mortality." *Studies in Family Planning* 23, 1: 23-33.

McCauley, A. P. and C. Salter. 1995. Meeting the Needs of Young Adults. *Population Reports,* Series J, No. 41. Baltimore, Johns Hopkins School of Public Health, Population Information Program.

McDonagh, M. 1996. "Is Antenatal Care Effective in Reducing Maternal Morbidity and Mortality?" *Health Policy Planning* 11, 1: 1-15.

McGinn, T. 2000, "Reproductive Health of War-Affected Populations: What Do We Know?" *International Family Planning Perspectives* 26, 4: 174-180.

MEASURE DHS+. 2000. *Service Provision Assessment (SPA) Tool Kit.* Baltimore, MD: Macro International Inc.

Measure DHS+. 2001. *Service Provision Assessment Guidelines and Field Instruments.* ORC MACRO, Calverton: MD. (Unpublished)

MEASURE *Evaluation* Project. 2001. *Quick Investigation of Quality (QIQ): A User's Guide for Monitoring Quality of Care in Family Planning.* MEASURE *Evaluation* Project Manual Series 2. University of North Carolina, Chapel Hill: Carolina Population Center.

Mensch, B. 2001. Personal communication.

Mensch, B., J. Bruce, and M. Greene. 1998. *The Uncharted Passage: Girls' Adolescence in the Developing World.* New York, NY: The Population Council.

Miller, R., A. Fisher, K. Miller, L. Ndhlovu, B. Ndugga Maggna, I. Askew, and D. Sanogo. 1997. *The Situation Analysis Approach to Assessing Family Planning and Reproductive Health Services: A Handbook.* New York, NY: The Population Council.

Ministry of Health [Kenya], National Council for Population and Development [Kenya], and ORC Macro. 2000. *Kenya Service Provision Assessment Survey 1999.* Calverton, Maryland: Ministry of Health, National Council for Population and Development, and ORC Macro.

Mitra S.N., A. Al-Sabir, A. Cross, and K. Jamil. 1997. *Bangladesh Demographic and Health Survey, 1996-1997.* Dhaka and Calverton, Maryland: National Institute of Population Research and Training (NIPORT), Mitra and Associates, and Macro International.

MotherCare. 2000a. "MotherCare's Community Assessments. Understanding Family and Community Behaviors and Practices." *Mothercare Matters* 8, 3-4: 1-17.

MotherCare. 2000b. "Behavioral Dimensions of Health and Survival." *Mothercare Matters* 9, 3: 1-19. MSH. 1999a *Management and Organizational Sustainability Tool.* Boston, MA: Management Sciences for Health.

MSH. 1999b. *Managing Drug Supply*. West Hartford, CT: Kumarian Press.

MSH. 1999c. *Understanding and Using Financial Management Systems to Make Decisions.* Boston, MA: Management Sciences for Health.

MSH. 2001. "Managing and Leading Framework." Management and Leadership Program (Unpublished)

Nazzar, A., L.L. Reason, P.B. Adongo, and J.F. Phillips. 2001. "A Community-informed Experiment in Preventing Female Genital Cutting among the Kassena-Nankana of Northern Ghana." Navrongo Health Research Centre, Ministry of Health, Box 114, Navrongo, Upper East Region, Ghana. (Unpublished)

Nelson, K., L. MacLaren, and R. Magnani. 2000. *Assessing and Planning for Youth-Friendly Reproductive Health Services*. FOCUS Tool Series 2. Washington, DC: FOCUS on Young Adults Program.

Nelson, K., R. Magnani, and K. Bond. 2001. *Does Making Health Services Youth-Friendly Increase Service Use by Adolescents? Evidence from Lusaka, Zambia*. Washington, DC: FOCUS on Young Adults Program.

NHS Centre for Reviews and Dissemination. 1999. *Getting Evidence into Practice* 5, 1. The University of York. Obermeyer, C. 1999. "Female Genital Surgeries: the Known, the Unknown, and the Unknowable." *Medical Anthropological Quarterly* 13, 1: 79-105.

Obermeyer, C. 1999. "Female Genital Surgeries: The Known, The Unknown, and The Unknowable." *Medical Anthropological Quarterly* 13, 1:79-105.

Otsea, K., J. Benson, D. Measham, M. Thorley, and R. Lidh. 1999. "Monitoring Postabortion Care." *Technical Resources for Postabortion Care* Vol. 3. Chapel Hill, NC: Ipas.

PAHO. 2001. "Integrated Vision for Vitamin A Supplementation in the Americas" Regional Meeting Report. Washington, D.C.

Pathak, L., B. Kwast, D.S. Malla, A.S.R.R. Pradham, and B. Campbell. 2001. *Process Indicators for Safe Motherhood Programmes: Their Application and Implications as Derived from Hospital Data in Nepal.* (Unpublished)

Perreira, K., P. Bailey, E. Bocalletti, and E.V.S. Hurtado. 2001. *Improving the Recognition of Obstetric Complications through Community and Clinic Based Education in Guatemala.*

Petitti, A.B., R.C. Cefalo, S. Shapiro, and P. Whalley. 1982. "In-hospital Maternal Mortality in the United States: Time Trends and Relation to Method of Delivery." *Obstetrics and Gynecology* 59, 1: 6-12.

Pilch, S.M. 1987. "Analysis of Vitamin A Data from the Health and Nutrition Examination Surveys." *Journal of Nutrition* 117, 4: 636-640.

Piotrow, P.T., D.L. Kincaid, M.J. Hindin, C.L. Lettenmaier, I. Kuseka, T. Silberman, A. Zinanga, F. Chikara, D.J. Adamchak, and M.T. Mbizvo. 1992. "Changing Men's Attitudes and Behavior: The Zimbabwe Male Motivation Project." *Studies in Family Planning* 23, 6.

POLICY Project. 1999. *Networking for Policy Change: An Advocacy Training Manual* I: Washington D.C.: The Futures Group International.

POLICY Project. 2001. *Planning Checklist.* Washington, D.C.: The Futures Group International.

Population Action International. 2001. *The PAI Report Card 2001 A World of Difference: Sexual and Reproductive Health and Risks*. Washington, DC: PAI.

The Population Council. 2000. *Postabortion Care Supervisory Aid and Evaluation Tool (PAC Guide)*. New York, NY: The Population Council.

Potter, J. E. 1999. "Outmoded Contraceptive Regimes: The Cases of Mexico and Brazil." *Population and Development Review* 25, 4: 709-739.

Price, D.D. 1983. "The Validation of Visual Analogue Scales as Ratio Scale Measures for Chronic and Experimental Pain." *Pain* 17: 45-56.

Price, D.D. and S.W. Harkins. 1987. "The Combined Use of Visual Analogue Scales and Experimental Pain in Providing Standardized Measurement of Clinical Pain." *Clinical Journal of Pain* 3: 1-8.

PriceWaterhouseCoopers. 2001. *The Performance Management Toolkit: A Guide to Developing and Implementing Performance Monitoring Plans.* Arlington, VA: PriceWaterhouseCoopers LLP.

PRIME. 2000. *PRIME performance improvement. L'Amelioration de la performance. Mejoramiento del Desempeno.* Program for International Training in Health [INTRAH]. Prime Project. Chapel Hill, North Carolina: INTRAH.

Prual, A. 2000. "Severe Maternal Morbidity from Direct Obstetric Causes in West Africa: Incidence and Case Fatality Rates." *Bulletin of the World Health Organization* 78, 5: 593-602.

QAP/URC. 2000a. "The Health Manager's Guide for Monitoring the Quality of Primary Care." Bethesda, MD: University Research Co., LLC.

QAP/URC. 2000b. "The Health Manager's Guide for Monitoring the Quality of In-Patient Care." Bethesda, MD: University Research Co., LLC.

QAP/URC. 2001a. "QA Monograph on Institutionalizing Quality Assurance." DRAFT. University Research Co., LLC.,

QAP/URC, 2001b. "A Modern Paradigm for Improving Quality of Health Care Services." Bethesda, MD: University Research Co, LLC.

QAP/URC. 2001c. "The Latin American Maternal Mortality (LAMM) Initiative." Bethesda, MD: University Research Co., LLC.

Reason, L. 2001. Personal communication.

Resnick, M., P. Bearman, R. Blum, K. Bauman, K. Harris, J. Jones, J. Tabor, T. Buering, R. Sieving, M. Shew, M. Ireland, L. Bearinger, and J. Udry. 1997. "Protecting Adolescents from Harm: Findings from the National Longitudinal Study on Adolescent Health." *Journal of the American Medical Association* 278, 10: 823-32.

Reyes J and L. Ochoa. 2001. Instituto Nacional de Estadística e Informática and ORC Macro International, Inc. 2001. *Perú: Encuesta Demográfica y de Salud Familiar* 2000. Calverton, Maryland: INEI and ORC Macro.

Robey, B., S.O. Rutstein, L. Morris, and R. Blackburn. 1992. "The Reproductive Revolution: New Survey Findings*." Population Reports*, Series M, 11. Johns Hopkins University. Baltimore, MD: Population Information Program.

Rogo, K.O., V.M. Lema, and G.O. Rae. 1999. *Postabortion Care: Policies and Standards for Delivering Services in Sub-Saharan Africa*. Chapel Hill, NC: Ipas.

Ronsmans, C., E. Achadi, G. Sutratikto, A. Zazri, and J. McDermott. 1999. "Use of Hospital Data for Safe Motherhood Programmes in South Kalimantan, Indonesia." *Tropical Medicine and International Health* 4, 7: 514-521.

Rooney, C. 1992. "Antenatal Care and Maternal Health: How Effective Is It? A Review of the Evidence." Division of Family Health, Maternal Health and Safe Motherhood Programme. Geneva: WHO.

Rosenfield A, 2001. "Maternal Mortality as a Human Rights and Gender Issue". In: *Reproductive Health, Gender and Human Rights: a dialogue.* Edited by Elaine Murphy and Karin Ringheim. Washington D.C.: Program for Appropriate Technology in Health [PATH].

Rosenfield A. and E. Figador. 2001. "Where is the M in MTCT? The Broader Issues in Mother to Child Transmission of HIV." *American Journal of Public Health* 91, 5: 703-4

Ross J.A., W.P. Mauldin, S.R. Green, and E.R Cooke. 1992. *Family Planning and Child Survival Program, as Assessed in 1991*. New York, NY: The Population Council

Ross J.A. and W.P. Mauldin. 1996. "Family Planning Programs: Efforts and Results." *Studies in Family Planning* 27, 3: 137-147.

Ross, J., O.M.R. Campbell and R. Bulatao. 1999. "The Maternal and Neonatal Programme Effort Index (MNPI)."*Journal of Tropical Medicine and International Health* 6,10: 787-798.

Ross, J. and J. Stover. 2001. "Effort Indices for National Family Planning Programs, 1999 Cycle." *International Family Planning Perspectives.* New York, NY: 27, 3: 119-129.

Rossi, P.H., H. Freeman, and M. Lipsey. 1999. *Evaluation: A Systematic Approach.* Thousand Oaks, CA: Sage Publications.

Rutstein, S.O. 1999. "Guidelines for the MEASURE *DHS*+ Main Survey Report." *Demographic and Health Surveys*. Baltimore, MD: Macro International Inc.

Sanday, P.R. 1981. "The Socio-Cultural Context of Rape: A Cross-Cultural Study." *Journal of Social Issues* 37, 4: 5-27.

Schuler, S.R. and Z. Hossain. 1998. "Family Planning Clinics Through Women's Eyes and Voices: A Case Study from Rural Bangladesh." *International Family Planning Perspectives* 24,4: 170-175, 205.

Schwarcz R., A. Diaz, and F. Nieto. 1990. "Guía para la vigilancia del parto. Partograma con curvas de alerta." Montevideo, Uruguay: Centro Latinoamericano de Perinatología.

Seiber, E. and J.T. Bertrand. 2001. "Access as a Factor in Differential Contraceptive Use between Mayans and Ladinos in Guatemala." MEASURE *Evaluation* Project Working Paper Series. University of North Carolina, Chapel Hill, NC: Carolina Population Center.

Senderowitz, J. 1999. "Making Reproductive Health Services Youth Friendly." *Research, Program and Policy Series*. Washington, DC: FOCUS on Young Adults Program.

Sen, A. 200. "Many Faces of Gender Inequality." *Frontline* 18, 22.

SIECUS, 1996. *Guidelines for Comprehensive Sexuality Education: Kindergarden through 12th grade.* 2nd Edition. New York, NY: Sexuality Information and Education Council of the United States.

Singh, S. and D. Wulf. 1994. "Estimated Levels of Induced Abortion in Six Latin American Countries." *International Family Planning Perspectives* 20, 1: 4-13.

Solo, J., A. Cerulli, R. Miller, I. Askew, and E. Pearlman. 1998. *Strengthening the Utilization of Family Planning Operations Research: Findings from Case Studies in Africa*. New York, NY: The Population Council.

The SPHERE Project. 2000. *Humanitarian Charter and Minimum Standards in Disaster Response*. Humanitarian Charter and Minimum Standards in Disaster Response, Geneva: The Sphere Project.

Stang, A., P. Schwingl, and R. Rivera. 2000. "New Contraceptive Eligibility Checklists for Provision of Combined Oral Contraceptives and Depot-Medroxyporgesterone Acetate in Community-Based Programmes." *Bulletin of the World Health Organization* 78, 8: 1015-1023.

Stanton, C., N. Abderrahim, and K. Hill. 2000. "An Assessment of DHS Maternal Mortality Indicators." *Studies in Family Planning* 31, 2: 111-123.

Stinson, W. 1995. Personal communication.

Stoltzfus, R.J. and B.A. Underwood. 1995. "Breast-milk Vitamin A as an Indicator of the Vitamin A Status of Women and Infants." *Bulletin of the World Health Organization* 5: 703-711.

Stover, J., J.T. Bertrand, and J. Shelton. 2000. "Empirically-based Conversion Factors for Calculating CYP." *Evaluation Review* 24, 1: 3-46.

Stover, J., B. Schwartlander, and J. Roehnstrom. 2000. "Measuring the Level of Effort in the National and International Response to HIV/AIDS: The AIDS Program Effort Index (API)." Paper presented at the XIII International AIDS Conference, July 9-14, Durban, South Africa.

Sullivan, T.M., J. Rice, and J.T. Bertrand, 2001. "Empirically Testing the Bruce/Jain Framework: Are Six Elements Enough?" Presented at Population Association of America Meeting, March 29-31, 2001, Washington, DC.

Sullivan, T.M. and J.T. Bertrand (eds). 2000. *Monitoring Quality of Care in Family Planning by the Quick Investigation of Quality (QIQ): Country Reports*. MEASURE *Evaluation* Project Technical Report Series 5. University of North Carolina, Chapel Hill, NC: Carolina Population Center.

Technical Guidance/Competence Working Group. 2001. *Recommendations for Updating Selected Practices in Contraceptive Use: Results of a Meeting: Volumes 1 and 2*. USAID. (Available online at: http://www.reproline.jhu.edu/english/6read/6multi/tgwg/6tgwg.htm)

Technical Guidance/Competence Working Group. 1997. *Recommendations for Updating Selected Practices in Contraceptive Use, Vol. II.* Chapel Hill, NC: INTRAH.

Thaddeus, S. and D. Maine. 1994. "Too Far to Walk: Maternal Mortality in Context." *Social Science Medicine* 38, 8: 1091-1110.

Toubia, N. 1995. *Female Genital Mutilation: A Call for Global Action*. New York, NY: Women, Inc.

Trostle J., M. Bronfman and A. Langer. 1999. "How Do Researchers Influence Decision-makers? Case studies of Mexican Policies." *Health Policy and Planning* 14, 2: 103-114.

Tsu, V. 2000. *Clean Home Delivery Kit: Evaluation of the Health Impact*. Seattle, WA: PATH.

Tsui, A.O., and L.H. Ochoa. 1992. "Service Proximity as a Determinant of Contraceptive Behavior: Evidence from Cross–national Studies of Survey Data." In J.F. Philips and J.A. Ross (eds) *Family Planning Programs and Fertilit*y. Oxford: Clarendon Press.

Tsui A, A. Hermali , J. Bertrand, J. Knowles, J. Stover, and K. Stewart. 1992. "Evaluating Family Planning Program Impact: New Initiatives on a Persisting Question." Presented at the Annual Meeting of the Population Association of America. Denver, CO. (Unpublished)

UNAIDS, 1998. Gender and HIV/AIDS: UNAIDS Technical Update. Geneva:UNAIDS.

UNAIDS, 1999a. Gender and HIV/AIDS: Taking Stock of Research and Programmes. Geneva: UNAIDS.

UNAIDS, 1999b. AIDS Five Years Since ICPD: Emerging Issues and Challenges for Women, Young People, and Infants. Geneva: UNAIDS.

UNAIDS. 2000. *National AIDS Programmes: A Guide to Monitoring and Evaluation.* Geneva: UNAIDS.

UNFPA. 2000. *Partners for Change: Enlisting Men in HIV/AIDS Prevention.* New York, NY: UNFPA.

UNFPA. 1998a. *Indicators for Population and Reproductive Health Programmes.* New York, NY: Technical and Policy Division.

UNFPA. 1998b. "Issues in Measuring and Monitoring Maternal Mortality. Implications for Programmes." *Technical and Policy Paper 1.* New York, NY: UNFPA.

UNFPA, UNICEF, The World Bank, WHO, IPPF, and The Population Council. 1997. *The Safe Motherhood Action Agenda: Priorities for the Next Decade.* Report on the Safe Motherhood Technical Consultation. October 18-23.

UNICEF, WHO, and UNFPA. 1997. *Guidelines for Monitoring the Availability and Use of Obstetric Services.* New York, NY: UNICEF.

United Nations. 1994. "Programme of Action of the International Conference on Population and Development in Cairo." Cairo: United Nations.

United Nations. 1995. *Population and Development: Programme of Action Adopted at the International Conference on Population and Development, Cairo, 5-13 September 1994.* Department for Economic and Social Information and Policy Analysis.

United Nations High Commissioner for Refugees. 1993. *Note on Certain Aspects of Sexual Violence against Refugee Women.* Geneva: UNHCR.

United Nations High Commissioner for Refugees. 1995. *Guidelines on Prevention and Response to Sexual Violence against Refugees.* Geneva: UNHCR.

United Nations High Commissioner for Refugees. 1999. *Reproductive Health in Refugee Situations: An Inter-Agency Field Manual.* Geneva: UNHCR.

USAID. 1998. *Handbook of Democracy and Governance Program Indicators:* 116-117. Center for Democracy and Governance. Washington D.C.: USAID.

USAID. 1999. *Health and Family Planning Indicators: Measuring Sustainability Volume II*. Africa Bureau, Office of Sustainable Development (AFR/SD). Washington D.C.: USAID.

USAID. 2000. "Monitoring the Policy Reform Process." *Recent Practices in Monitoring and Evaluation TIPS* 14. Center for Development Information and Evaluation. Washington D.C.: USAID.

Villar, J. and D. Khan-Neelofur. 2000. "Patterns of Routine Antenatal Care for Low-risk Pregnancy." *Cochrane Database Systems Review* 2: CD000934.

Walt, G. and L. Gilson. 1994. "Reforming the Health Sector in Developing Countries: The Central Role of Policy Analysis." *Health Policy and Planning* 9: 353-370.

West, K.P., J. Katz, S.B. Khatry, S.C. LeClerq, E.K. Phadhan, S.R. Shrestha, P.G. Connor, S.M. Dali, P. Christian, R.P. Pokhrel, and A. Sommer. 1999. "Double Blind, Cluster Randomized Trial of Low Dose Supplementation with Vitamin A or B-carotene on Mortality Related to Pregnancy in Nepal." *British Medical Journal* 318: 570-575.

Westoff, C.F. 1991. *Reproductive Preference*s. Demographic and Health Surveys Comparative Studies 3. Columbia, MD: Institute for Resource Development/Macro International Inc.

Westoff, C.F. and L.H. Ochoa. 1991. *Unmet Need and Demand for Family Plannin*g. Demographic and Health Surveys Comparative Studies 5. Columbia, MD: Institute for Resource Development/Macro International Inc.

Westoff CF and A. Bankole, 2001. *The Contraception-Fertility Link in Sub-Saharan Africa and in Other Developing Countries.* DHS Analytical Studies No. 4. Calverton, MD: ORC Macro, MEASURE DHS+.

Whitney, E.N., C. Cataldo, and S.R. Rolfes. 1998. *Understanding Normal and Clinical Nutrition*. Belmont, CA: Wasdworth Publishing Company.

Williams, T., J. Schutt-Aine, and Y. Cuca. 2000. "Measuring Family Planning Service Quality Through Client Satisfaction Exit Interviews*." International Family Planning Perspectives*. 26(2)

Wishik, S.M. and K. Chen. 1973. *Couple–years of Protection: A Measure of Family Planning Program Output*. New York, NY: Columbia University, International Institute for the Study of Human Reproduction.

Witte, K. 1992. "Putting the Fear back into Fear Appeals: The Extended Parallel Process Model." *Communication Monographs*, 59: 329 349.

Witte, K. 1998. "Fear as Motivator, Fear as Inhibitor: Using the Extended Parallel Process Model to Explain Fear Appeal Successes and Failures." In P. A. Andersen and L. K. Guerrero (eds) *The Handbook of Communication and Emotion: Research, Theory, Applications, and Contexts*: 423-450. San Diego, CA: Academic Press.

WHO. 1977. "International Classification of Disease, 1975 Revision". World Health Organization 1:764. Geneva: WHO.

WHO. 1987. *Studying Maternal Mortality in Developing Countries: A Guidebook.* WHO/FHE/87.7

WHO. 1991a. *Essential Elements of Obstetrical Care at First Referral Level.* Geneva: WHO.

WHO. 1991b. *Maternal and Perinatal Infections: A Practical Guide. Report of a WHO Consultation.* Maternal Health and Safe Motherhood Programme, Division of Family Health. Geneva: WHO.

WHO. 1991c. *Indicators for Assessing Breast-feeding Practices*. Report of an Informal Meeting. June.

WHO. 1992. "International Statistical Classification of Diseases and Related Health Problems, 10th Revision." *WHO 1:1238.* Geneva: WHO.

WHO. 1993. "Global Prevalence of Iodine Deficiency Disorder." *Micronutrient Deficiency Information System Working Paper 1.*Geneva: WHO/UNICEF/ICCIDD.

WHO. 1994a. *Indicators to Monitor Maternal Health Goals: Report of a Technical Working Group 8-12 November 1993.* Geneva: WHO.

WHO. 1994b. "Mother-Baby Package: Implementing Safe Motherhood in Countries." *MotherCare Matters* 8, 1.

WHO. 1995a. *Maternal Anthropometry and Pregnancy Outcomes: A WHO Collaborative Study*. Geneva: WHO.

WHO. 1995b. *Complications of Abortion, Technical and Managerial Guidelines for Prevention and Treatment.* Geneva: WHO.

WHO. 1996a. *Mother-Baby Package: Implementing Safe Motherhood in Countries*. Maternal Health and Safe Motherhood Program, Division of Family Health. WHO/FHE/MSM/94.11. Geneva: WHO.

WHO. 1996b. *Essential Newborn Care: Report of a Technical Working Group, April 1994.* Maternal and Newborn Health/Safe Motherhood Unit. Division of Reproductive Health (Technical Support). Geneva: WHO.

WHO. 1996c. "Indicators for Assessing Vitamin A Deficiency and Their Application in Monitoring and Evaluating Intervention Programmes." *WHO Nutrition* 96.10. Geneva: WHO.

WHO. 1996d. *Female Genital Mutilation: Report of a Technical Working Group.* Geneva: WHO.

WHO. 1997. *Monitoring Reproductive Health: Selecting a Short List of National and Global Indicators.* Family and Reproductive Health. Geneva: WHO.

WHO. 1998a. *Postpartum Care for the Mother and Newborn: a Practical Guide*. Maternal and Newborn Health/ Safe Motherhood Unit. Division of Reproductive Health (Technical Support). Geneva: WHO/RHT/MSM/98.3

WHO. 1998b. *WHO Recommended Standards for Surveillance of Selected Vaccine-preventable Diseases*. Global Programme for Vaccines and Immunisation. Expanded Programme on Immunisation. Geneva: WHO.

WHO. 1998c. *Unsafe Abortion: Global and Regional Estimates of Incidence of and Mortality due to Unsafe Abortion with a Listing of Available Country Data,* Third Edition. Geneva: WHO.

WHO. 1999a. *Indicators to Monitor Maternal Health Goals: Report of a Technical Working Group, Geneva, 8-12 November 1993.* Division of Family Health. Geneva: WHO.

WHO. 1999b. *Reduction of Maternal Mortality, a Joint WHO, UNFPA, UNICEF, World Bank Statement.* Geneva: WHO.

WHO. 1999c. *Field Manual for Neonatal Tetanus Elimination.* Department of Vaccines and Other Biologicals. WHO/V&B/99.14. Geneva: WHO.

WHO. 1999d. *Putting Women First: Ethical and Safety Recommendations for Research on Domestic Violence against Women.* Geneva. (Unpublished)

WHO. 2000a. "Reproductive Health Indicators for Global Monitoring: Report of the Second Interagency Meeting, Geneva 17-19th July 2000." WHO/RHR/01. Deparment of Reproductive Health and Research, Family and Community Health. Geneva: WHO.

WHO. 2000b. *The Management of Nutrition in Major Emergencies*. ISBN 92 4 154520 8. Geneva: WHO.

WHO. 2001a. "Fifty-fourth World Health Assembly": May 14-22, 2001. Geneva: WHO.

WHO. 2001b. *Essential Care Practice Guide: Pregnancy, Childbirth and Newborn Care*. Integrated Management of Pregnancy and Childbirth. Geneva. (Unpublished)

WHO. 2001c. *Making Pregnancy Safer: A Health Sector Strategy for Reducing Maternal and Perinatal Morbidity and Mortality*. Department of Reproductive Health and Research. Geneva: WHO.

WHO/MI. 1998. *Safe Vitamin A Dosage During Pregnancy and Lactation: Recommendations and Report of a Consultation*. Geneva: WHO.

WHO, UNICEF, and UNFPA. 2001. *Maternal Mortality in 1995: Estimates Developed by WHO, UNICEF and UNFPA*. WHO/RHR/01.9. Geneva: WHO.

WHO, UNICEF, and Wellstart International. 1999. *The Baby Friendly Hospital Initiative: Monitoring and Reassessment: Tools to Sustain Progress*. Geneva: WHO.

World Bank. 2001. *Engendering Development: Through Gender Equality in Rights, Resources, and Voice*. Policy Research Report. Washington, DC: The World Bank.

Wulf, D. 1994. *Refugee Women and Reproductive Health Care: Reassessing Priorities*. Women's Commission for Refugee Women and Children. New York, NY.

Yinger, N., A. Peterson, M. Avni, J. Gay, R. Firestone, K. Hardee, E. Murphy, B. Herstad, and C. Johnson-Welch. 2001. *Mainstreaming Gender in Monitoring and Evaluation: A Practical Approach for Reproductive Health and Nutrition Programming*. Interagency Gender Working Group, Subcommittee on Research and Indicators. (Unpublished)

Yoder, S., P.O. Camara, and B. Soumaoro. 1999. *Female Genital Cutting and Coming of Age in Guinea.* Calverton, MD: MACRO International Inc. and Conakry, Guinea: Universite de Conakry.